(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B6

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
9 August 2001 (09.08.2001)

PCT

(10) International Publication Number
WO 01/57190 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US01/04098

(22) International Filing Date: 5 February 2001 (05.02.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

09/496,914	3 February 2000 (03.02.2000)	US
09/560,875	27 April 2000 (27.04.2000)	US
09/598,075	20 June 2000 (20.06.2000)	US
09/620,325	19 July 2000 (19.07.2000)	US
09/654,936	1 September 2000 (01.09.2000)	US
09/663,561	15 September 2000 (15.09.2000)	US
09/693,325	20 October 2000 (20.10.2000)	US
09/728,422	30 November 2000 (30.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	09/496,914 (CIP)
Filed on	3 February 2000 (03.02.2000)
US	09/560,875 (CIP)
Filed on	27 April 2000 (27.04.2000)
US	09/598,075 (CIP)
Filed on	20 June 2000 (20.06.2000)
US	09/620,325 (CIP)
Filed on	19 July 2000 (19.07.2000)
US	09/654,936 (CIP)
Filed on	1 September 2000 (01.09.2000)
US	09/663,561 (CIP)
Filed on	15 September 2000 (15.09.2000)
US	09/693,325 (CIP)
Filed on	20 October 2000 (20.10.2000)
US	09/728,422 (CIP)
Filed on	30 November 2000 (30.11.2000)

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **LIU, Chenghua** [CN/US]; 1125 Ranchero Way, Apt. #14, San Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 Greenwich Place, Palo Alto, CA 94303 (US). **ASUNDI, Vinod** [US/US]; 709 Foster City Boulevard, Foster City, CA 94404 (US). **ZHOU, Ping** [CN/US]; 7595 Newcastle Drive, Cupertino, CA 95014 (US). **XU, Chongjun** [US/US]; 4918 Manitoba Drive, San Jose, CA 95130 (US). **CAO, Yicheng** [CN/US]; 260 N. Mathilda Avenue, #E6, Sunnyvale, CA 94086 (US). **MA, Yunquing** [CN/US]; 280 W. California Avenue, #206, Sunnyvale, CA 94086 (US). **ZHAO, Qing, A.** [CN/US]; 1556 Kooser Road, San Jose, CA 95118 (US). **WANG, Dunrui** [CN/US]; 932 La Palma Place, Milpitas, CA 95035 (US). **WANG, Jian-Rui** [CN/US]; 744 Stendahl Lane, Cupertino, CA 95014 (US). **ZHANG, Jie** [CN/US]; 4930 Poplar Terrace, Campbell, CA 95008 (US). **REN, Feiyan** [US/US]; 7703 Oak Meadow Court, Cupertino, CA 95014 (US). **CHEN, Rui-hong** [US/US]; 1031 Flying Fish Street, Foster City, CA 94404 (US). **WANG, Zhi, Wei** [CN/US]; 836 Alturas Avenue, #B6, Sunnyvale, CA 94085 (US). **XUE, Aidong, J.** [CN/US]; 1621 S. Mary Avenue, Sunnyvale, CA 94087 (US). **YANG, Yonghong** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **WEJHRMAN, Tom** [US/US]; CCSR Mol Pharm 3210, 269 W. Campus Drive, Stanford, CA 94305 (US). **GOODRICH, Ryle** [US/US]; 4896 Sandy Lane, San Jose, CA 95124 (US).

(74) Agent: **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky, and Popeo, P.C., One Financial Center, Boston, MA 02111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) **Title:** NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

(57) **Abstract:** The present invention provides novel nucleic acids, novel polypeptide sequences encoded by these nucleic acids and uses thereof.

WO 01/57190 A2

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954. The polypeptides sequences are designated SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954. The sequence information can be a segment of any one of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954 that uniquely identifies or represents the sequence information of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety

of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readablemedia, use in sequencing full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954 ; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the

invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, *e.g., in situ* hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (*e.g.*, bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting

symptoms or tendencies. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Tables 2 and 9); for which they have a signature region (as set forth in Tables 3 and 10); or for which they have homology to a gene family (as set forth in Tables 4 and 11). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

It must be noted that as used herein and in the appended claims, the singular forms “a”, “an” and “the” include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms “biologically active” or “biological activity” refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise “immunologically active” or “immunological activity” refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells” as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms “complementary” or “complementarity” refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5’-AGT-3’ binds to the complementary sequence 3’-TCA-5’. Complementarity between two single-stranded molecules may be “partial” such that only some of the nucleic acids bind or it may be “complete” such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides, more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100

nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides, preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-20.

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954. The sequence information can be a segment of any one of SEQ ID NO:1-1-984, 1969-2952, 3937-3942 or 3949-3954 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4^{20} possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match ($1 \div 4^{25}$) times the

increased probability for mismatch at each nucleotide position (3 x 25). The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms “operably linked” or “operably associated” refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term “pluripotent” refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

The terms “polypeptide” or “peptide” or “amino acid sequence” refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 500 amino acids, more preferably less than 200 amino acids more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term “translated protein coding portion” means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term “mature protein coding sequence” means a sequence which encodes a peptide or protein without a signal or leader sequence. The “mature protein portion” means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue

may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant"(or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, *e g.*, recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making

insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, *e.g.*, polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (*e.g.*, microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, *e.g., E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can

comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (*e.g.*, soluble proteins) or partially (*e.g.*, receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization

to filter-bound DNA in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (*i.e.*, the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, *e.g.*, mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, *e.g.*, mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% sequence identity, more preferably at least 98% sequence identity and most preferably at least 98% idenity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% identity, more preferably at least about 85% identity, more preferably at least about 90% identity, and most preferably at least about 95% identity, more preferably at least 98% and most preferably at least about 99% identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of

determining equivalence, truncation of the mature sequence (*e.g.*, via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J. (1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing as SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960; (c) a

polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO:985-1968, 2953-3936, 3943-3948 or 3955-3960. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, *e.g.*, at least about 65%, at least about 70%, at least about

75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, and more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, *e.g.*, by substituting first with conservative choices (*e.g.*, hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (*e.g.*, hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., *DNA* 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, *Nucleic Acids Res.* 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired

amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., *Gene* 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and *Current Protocols in Molecular Biology*, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression

vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct

transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

4.3 ANTISENSE

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding a nucleic acid disclosed herein (*e.g.*, SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the

strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res* 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res* 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue *et al.* (1987) *FEBS Lett* 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be designed based upon the nucleotide sequence of a DNA disclosed herein (*i.e.*, SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Pat. No. 4,987,071; and Cech *et al.* U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel *et al.*, (1993) *Science* 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (*e.g.*, promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991) *Anticancer Drug Des.* 6: 569-84; Helene. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup *et al.* (1996) *Bioorg Med Chem* 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup *et al.* (1996) above; Perry-O'Keefe *et al.* (1996) *PNAS* 93: 14670-675.

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, *e.g.*, inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, *e.g.*, in the analysis of single base pair mutations in a gene by, *e.g.*, PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, *e.g.*, S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup *et al.* (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, *e.g.*, to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, *e.g.*, RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn *et al.* (1996) *Nucl Acids Res* 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag *et al.* (1989) *Nucl Acid Res* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn *et al.* (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen *et al.* (1975) *Bioorg Med Chem Lett* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.* 86:6553-6556; Lemaitre *et al.*, 1987, Proc. *Natl. Acad. Sci.* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization triggered cleavage agents (See, *e.g.*, Krol *et al.*, 1988, *BioTechniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon, 1988, *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a

peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis.*

The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5′ flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or

glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.*, inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or (b) polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, and more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (*e.g.*, an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, *e.g.*, Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: *A Laboratory Manual*; Ausubel et al., *Current Protocols in Molecular Biology*. Polypeptide fragments that retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for e.g., small molecules, molecules from combinatorial libraries, antibodies or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960.

The protein of the invention may also be expressed as a product of transgenic animals, *e.g.,* as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, *e.g.*, U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, Calif., U.S.A. (the MaxBat™ kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (*i.e.*, from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl™ or Cibacrom blue 3GA Sepharose™; one or more steps involving

hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP- HPLC) steps employing hydrophobic RP-HPLC media, *e.g.*, silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, e.g., antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al. NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.
In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprise one or more domains fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and

administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e,g.,* cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected *ex vivo, in situ,* or *in vivo* by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of

the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered *in vivo* to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may

be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, *e.g.,* inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to

identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse

and human interleukin-γ, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6--Nordan, R. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9--Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, *In Vitro* assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells *in vivo* or *ex vivo* is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of

cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention, optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation

of neural cells and for the regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support e.g. as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e.,

traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book

Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof, including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (*e.g.*, anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastbom et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization

test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain

protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery

et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma, acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including

bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine. Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These *in vitro* models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissue Culture Collection catalogs.

4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley- Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1- 7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or “leads” via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for

screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science 282*:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., *Mol. Biotechnol*, 9(3):205-23 (1998); Hruby et al., *Curr Opin Chem Biol*, 1(1):114-19 (1997); Dorner et al., *Bioorg Med Chem*, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules, that modulate (*i.e.,* increase or decrease) biological activity of a polypeptide of the invention.

Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications i.e. phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid

arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, *etc.*, depending on the molecule to be measured; and motor neuron dysfunction may be measured by

assessing the physical manifestation of motor neuron disorder, *e.g.*, weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a

suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01μg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1μg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, *e.g.*, treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral

ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the

pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic,

talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, *e.g.*, gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated

solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium

carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present

invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns.

In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.,* amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a

mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, *e.g.*, Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the

desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about 0.01 μg/kg to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 μg/kg to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab}, $F_{ab'}$ and $F_{(ab')2}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well,

such as IgG_1, IgG_2, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO:985, and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory

Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

5.13.2 Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen

binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro. The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the

Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal. The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', $F(ab')_2$ or other antigen-binding subsequences of antibodies) that are principally comprised of the sequence of a human

immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach

is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al,(Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the Xenomouse™ as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another

mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab')2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab')2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion

preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody $F(ab')_2$ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).
Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins

can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi, ^{131}I, ^{131}In, ^{90}Y, and ^{186}Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido

compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring

formats (*e.g.* text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing

software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic

acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide *in vivo* at the target site.

4.18 SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:

1-984, 1969-2952, 3937-3942 or 3949-3954, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed.

As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

4.19 USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The

hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent *in situ* hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

Oligonucleotides, i.e., small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.,* 1985; Dahlen *et al.,* 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.,* 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, e.g., Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen *et al.,* (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm$_7$), is then added to a final concentration of 10 mM 1-MeIm$_7$. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm$_7$, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook *et al.* (1989), shearing by ultrasound and NaOH treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *Cvi*JI, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease *Cvi*JI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (*Cvi*JI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald *et al.* (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a *Cvi*JI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that *Cvi*JI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5

ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm^2, depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane. Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm^2 and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and

variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

Assemblage of Novel Nucleic Acids

The contigs or nucleic acids of the present invention, designated as SEQ ID NO: 1969-2951, and 3949-3954 were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Tables 6 and 8 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:2953-3936, and 3949-3954) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 2953-3936 and 3955-3960. Tables

6 and 8 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

5.3 EXAMPLE 3

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), full length gene cDNA sequences and their corresponding protein sequences were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genebank. Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide sequences are shown in the Sequence Listing as SEQ ID NO:1-351. The amino acids are SEQ ID NO:985-1335.

Table 1 shows the various tissue sources of SEQ ID NO: 1-351.

The nearest neighbor results for SEQ ID NO: 1-351 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 1-351 from Genpept . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 1-351 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.4 EXAMPLE 4

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e. dbEST version 117, gb pri 117, UniGene version 117, Genpept release 117). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 352-766. The corresponding amino acids are SEQ ID NO: 1336-1750.

Table 1 shows the various tissue sources of SEQ ID NO: 352-766.

The nearest neighbor results for SEQ ID NO: 352-766 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 352-766 from Genpept . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 352-766 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.5 EXAMPLE 5

__Novel Nucleic Acids__

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e., dbEST version 118, gb pri 118, UniGene version 118, Genpept release 118). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 767-930. The corresponding amino acid sequences are SEQ ID NO:1751-1914.

Table 1 shows the various tissue sources of SEQ ID NO: 767-930.

The homology results for SEQ ID NO: 767-930 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release 21(Derwent), using BLAST algorithm. The nearest neighbor result showed the homologs for SEQ ID NO: 767-930 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologues with identifiable functions for SEQ ID NO: 767-930 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.6 EXAMPLE 6

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e. dbEST version 118, gb pri 118, UniGene version 118, Genpept release 118). Other computer programs which may have been used

in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS: 931-965. The corresponding amino acid sequences are shown in SEQ ID NO:1915-1949.

Table 1 shows the various tissue sources of SEQ ID NO: 931-965.

The nearest neighbor results for SEQ ID NO: 931-965 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 931-965 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 931-965 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.7 EXAMPLE 7

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e. dbEST version 119, gb pri 119, UniGene version 119, Genpept release 119). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS:966-974. The corresponding amino acid sequences are SEQ ID NO:1950-1958.

Table 1 shows the various tissue sources of SEQ ID NO: 966-974.

The nearest neighbor results for SEQ ID NO: 966-974 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 966-974 from Genpept . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 966-974 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in

each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.8 EXAMPLE 8

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e. dbEST version 120, gb pri 120, UniGene version 120, Genpept release 120). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS:975-984. The corresponding amino acid sequences are SEQ ID NO:1959-1968.

Table 1 shows the various tissue sources of SEQ ID NO: 975-984.

The nearest neighbor results for SEQ ID NO: 975-984 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 21, 2000 release (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 975-984 from Genpept. The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 975-984 are shown in Table 2 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 3 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 4 shows the name of the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also

disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 7 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

5.9 EXAMPLE 9

Novel Nucleic Acids

Using PHRAP (Univ. of Washington) or CAP4 (Paracel), a full length gene cDNA sequence and its corresponding protein sequence were generated from the assemblage. Any frame shifts and incorrect stop codons were corrected by hand editing. During editing, the sequence was checked using FASTY and/or BLAST against Genbank (i.e. dbEST version 120, gb pri 120, UniGene version 120, Genpept release 120). Other computer programs which may have been used in the editing process were phredPhrap and Consed (University of Washington) and ed-ready, ed-ext and gc-zip-2 (Hyseq, Inc.). The full-length nucleotide, including splice variants resulting from these procedures are shown in the Sequence Listing as SEQ ID NOS:3937-3942. The corresponding peptide sequence is SEQ ID NO: 3943-3948.

Table 1 shows the various tissue sources of SEQ ID NO: 3937-3942.

The nearest neighbor results for SEQ ID NO: 3937-3942 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq October 12, 2000 release 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the closest homologue for SEQ ID NO: 3937-3942 from Genpept . The translated amino acid sequences for which the nucleic acid sequence encodes are shown in the Sequence Listing. The homologs with identifiable functions for SEQ ID NO: 3937-3942 are shown in Table 9 below.

Using eMatrix software package (Stanford University, Stanford, CA) (Wu et al., J. Comp. Biol., Vol. 6 pp. 219-235 (1999) herein incorporated by reference), all the sequences were examined to determine whether they had identifiable signature regions. Table 10 shows the signature region found in the indicated polypeptide sequences, the description of the signature, the eMatrix p-value(s) and the position(s) of the signature within the polypeptide sequence.

Using the pFam software program (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1) pp. 320-322 (1998) herein incorporated by reference) all the polypeptide sequences were examined for domains with homology to certain peptide domains. Table 11 shows the name of

the domain found, the description, the p-value and the pFam score for the identified domain within the sequence.

The nucleotide sequence within the sequences that codes for signal peptide sequences and their cleavage sites can be determine from using Neural Network SignalP V1.1 program (from Center for Biological Sequence Analysis, The Technical University of Denmark). The process for identifying prokaryotic and eukaryotic signal peptides and their cleavage sites are also disclosed by Henrik Nielson, Jacob Engelbrecht, Soren Brunak, and Gunnar von Heijne in the publication " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering, Vol. 10, no. 1, pp. 1-6 (1997), incorporated herein by reference. A maximum S score and a mean S score, as described in the Nielson et as reference, was obtained for the polypeptide sequences. Table 12 shows the position of the signal peptide in each of the polypeptides and the maximum score and mean score associated with that signal peptide.

Tables 5 and 13 are correlation tables of all of the sequences and the SEQ ID NOS.

TABLE 1

Tissue Origin	RNA Source	Library Name	SEQ ID NOS:
lung			3 11 25 49 65 75 114 141 156 160 172 190 198 209 217 224 229 234-235 267 269 274 277 282 284 303 308 312 320 334 336 352 372 396 398 412 414 437 453 464 470 481 492-494 508-509 532 539 581 584 617-619 621 628 633 643 688 691 745 752 761 768 794 822 837 848 876 887 953 967 973
adult brain	GIBCO	AB3001	1 3 12-13 16 22-24 28-29 41 48 58 65 78 82 89-90 94 97 103 112 114-115 117 120 122 130-131 168 181 184 186-187 189-190 198 208 216 247 249 259 270 277 297 301 308 312 314 321 333 348 374 396 403 406 410 412 416-417 420 423 426-427 431 456 474 481 484-485 488 498 500 508-509 530 549 553 558 563-564 583 596 602-603 608 612 621-622 624 643 650 674 699 711 736 738-739 753 770 779-780 785-786 802-803 816 822 839 842 848 859 861 871 893-894 897 900 903 925 954 958 967 969
adult brain	GIBCO	ABD003	3 19 21-25 28-29 31 33-34 37 39 41 46-48 53 58 63-64 66 72 78 80 99 103 109-110 112 114 118 120-124 126 132-133 135

			139 143 146 148-149 159 163 168 174 176 179-180 184-185 188-190 202 208-209 216-217 221 223 230 234-235 240 244 249 251 253 255 258-259 263 269-270 277 282 285-286 290 294-295 297 301-302 304-305 307-308 311-312 314 320 329 333 335-336 342 344 346 349 354 358 365 370 373-374 377 380 382-383 388 394-396 399 401-402 406 409-410 413 416 420-421 425 428 430-431 436-437 442 456 462 464 466-467 474 484 486 495-496 500-501 506 508-509 519 530 537 542 549 561-562 564 572 574 577-578 580-583 586-587 589 592-593 596-597 601 608 610 612-614 617-624 630-632 635 637 650 658 663-664 668 676 679 681 689-690 693 699 724 726 732 736 742-743 747 767-770 780 784 789 793 799 802-805 813 817-818 822 824 829-831 837 839 845 848 856 859-860 864 871-872 875-876 881 887 896-897 901 903 907 910-911 925 930 933 943-944 947 952-953 958 962-963 965 967 972 977
adult brain	Clontech	ABR001	3 53 66 113 115 126 135 160 172 179 185 204 263 273 305 312 323 358 380 383 395-396 403 420 428-429 431 461 542 583 586 606-607 611 620 645-646 688 690 715 732 736 740 748 754 768 784-786 790 796 800 878 897 906-907 947 977
adult brain	Clontech	ABR006	19 32 49 53 60 72 91 103 118 125 130-131 134 184 224 275 338 350 354 361-363 374 384 390 394 396 431-432 434-435 445 468 549 621 732 734-736 745 760-761 764 768-769 775 787 806 811 818 887 903 906 918 930 942 947 957 973 977
adult brain	Clontech	ABR008	2-3 9-11 14 17 21 23-25 28-29 31-35 37 41-42 45 47-48 56-57 65-66 69-70 72 75 77-78 88 91-92 97-99 101 103 112-115 118-128 130-131 135 138-140 142 144-146 148 152 156-157 159-160 163 168 172 174 176 178-180 182-190 194 196-198 200-201 204 209-214 218 220-225 228-230 232-233 238-240 243-244 246 254-256 260-264 270 272-274 278-279 282-285 289-291 293-294 296-297 301 303-306 312-314 317 321-322 325-328 334 336 338 340-342 344 346 348 350-352 354 356-358 363 366 369-374 376 379-381 383-386 388-394 398-399 402-

			403 405 409-412 414 418-421 423-424 426-427 430 433-437 443 445-450 452 456-457 460 462 464 471 479 482-483 485 488 490-498 505 507 510 516 519-522 524 527-532 535 538-539 542-545 548 551 553 555 561-562 566 569 571 574 580-583 588-589 593 597 601-608 611-612 614-615 617-618 621-622 624 630-635 642 644 646-648 650-652 655 657 659-661 664-665 668 672 674 689 693-699 701-702 708 711 715 717 724 728-730 732 734-735 738-740 745 747-750 753-755 757 761 763-764 766-769 772-773 775 780-781 789-791 793-795 799-800 802-806 809 812 818-819 821-822 826 829-830 832 834-835 841 843 845 856 858-859 861 864 866 870 872 876 880 883 885 887 893-898 902 906-916 918 921 925-926 930-931 933 942-943 946 948 950-951 953-954 958-960 962-965 967 969-970 972 977
adult brain	Clontech	ABR011	57 196 270 304 344 436 834
adult brain	BioChain	ABR012	14 82 121-122 168 691
adult brain	Invitrogen	ABR013	72 108 263 270 336 425 492-494 732 787 790 826 880
adult brain	Invitrogen	ABR014	293 394 399 764 768-769 928 967
adult brain	Invitrogen	ABR015	738-739 764
adult brain	Invitrogen	ABR016	320 374 396 399 405 684 742-743 767 931 947 967
adult brain	Invitrogen	ABT004	21 33-34 37-38 47 52 57-58 69 72 91-93 109 119 122-124 126-127 135 142-143 158 167-168 185-188 194 200 212 232 242 246 255 258 270 277 279 293 301 312-313 319 322-323 331 341 346 348 371 374 388 391 394 399 401 409 411 429 436-437 456 462 477 488 496 498 510 512 515 539 542 545 549 559 563 573 579 587 589 601-605 612 620-621 624 640 643 647 681 715 723 728 732 735-736 740 745 748 753 766 785-786 792-793 797-801 812 822 829-831 853-856 859 876-877 884 893-894 908-909 918 925 933 950 969 978
cultured preadipocytes	Strategene	ADP001	4 28-29 69 93 114 121 132-133 135 151-152 159 167 172 178 181 184 190 194-195 203-204 209 217 219 240 248 260-262 267 273-274 277 282 297 301 304 312 314 326-327 361-362 371 374 388 394 401 403 405 411 420 437 453 466-467 470 474 478 496 507-509 517 530 532-533 584 588 593 602-603 608 610 617-621 630-631 633 639 642-643 661

			693 729 746 761 765 769 834 842 848 887 907 923 947-950 957 967 969
adrenal gland	Clontech	ADR002	1 3 12-13 21 23-24 27-29 67 74 78 103-105 108-109 113 115 118 120-121 128-133 149 156 160 172 177 182 214 217 223 232-233 247 254 269-270 273-274 277 283 285 288 298-299 308 317 319 328 338 340 342 361-362 364 372 376-377 382 384 401-402 405-406 416 420 431 437 444 446 448 457 462 484 500 507 517 524 532-533 539 545 554 561-562 564 588 597 602-603 606-607 635 642 646 649 658 664 674 693 703 730 740 745 752 759 765 767 775 779 799 809 817-818 839 845 856 859 863 887 890-891 896 948 953 958 961-963 973
adult heart	GIBCO	AHR001	1 3-4 8 10 14 20-21 25 28-29 33-34 37-38 41 48 54-57 65 69-72 75 78 80 82-83 97 99-100 108 112-115 117-121 123-124 128-133 141 144-146 149 152 159 162-163 168 172 176 179 181 184 186-187 190-191 201 203 208-209 212 216-218 221 223 227 229 233 244 247 249 253-255 258 263-264 267 269-270 274 278 280-282 285 289 291 295 297-299 301 303-304 308 313 317 321-322 326 328 334 344 348 352 358 361-363 370-371 380 382-383 388 394-396 398 401 403 405-406 410-416 423 425-427 430-431 436 452-453 464-465 470-474 481-484 487-488 490 492-494 496 499-500 505-506 508-509 514 523 529-530 533 547-548 553 558 563-565 577-578 586-588 590 593 597 601-603 606-608 610-613 617-619 621-622 626-628 637-638 642-644 652 658 661 672 682-683 688 691 693 697 699 708 711 713 715 732 737 745 747-748 750-753 759 761 765 768-770 775 790 802-803 814-815 818-819 830 837 839-840 842 845 848 859 861-862 867 876-877 887 891-892 896 900-901 903 905-906 908-909 919-920 922 925 928 936 939-940 946-947 950 953 959 967 970-971 973 977
adult kidney	GIBCO	AKD001	1 3 8 12-14 17 19-25 28-29 33-34 37-39 41 46-48 50 52 55-60 62 65-67 69 71-72 75 77-78 82 84 89-90 93 97 108-110 114-116 118-121 123-125 128 130-133 135 138 144 146 149 156 159-161 163-164 167-172 176 179 184 186-187 189-190 194 196 200-202 204 209 211-212 216-217 219 221 223-224 229 232-235 244

			247 250 253 255-256 258 263-264 268-272 274 277-281 283 286 288-290 292 294-295 297 301 303-309 311-314 316 319-323 325 328-338 342 348-349 352 354-355 358 361-363 365 370-371 373 376-378 380 382-383 388 395-399 401-403 405-406 409-413 416 418-420 425-428 430-431 440 442 452-454 462 464-465 470 472-474 477 479 481 483-485 487-489 492-495 498-500 504 506 510 517 522 525 529-530 532-533 539 542-543 547 551-552 558 560-564 569-570 573-574 577-578 580-583 585-590 594-596 601-608 610-613 617-621 624 626-628 630-631 634-636 639 642-643 648 652 656 658 664-665 676-677 679 681 688-691 693 697 699 708 711 715 717 720-722 724 729-732 738-741 747-748 751-753 761 765 770-778 780 784 789 791 793 797 804 813 817 823-824 834 837 839 842-843 845 848 859 861-862 864 867 870 876-877 887 889 892-894 896-897 900-901 903 907 913-915 918 921 923 925 929-930 932 939 942 946-947 949-950 953 958-959 961-963 967 969 972 977
adult kidney	Invitrogen	AKT002	1 3 16 21 30 32 35 38-41 46-47 56 77 92 109 123-124 130-131 146 149 161 167-168 172 176 190 209 212 234-235 258 279 292 301 303 308 314 333 355 363 372 380 383 396 399 402 418-419 426-427 431 448 454 461 471-474 488-489 495 498 504 506 508-509 520-521 530 537 539-541 545 547 563 582-583 592 613 617-618 621 623-624 633 655 688 690 693 699 704 713 732 745 752-753 761 766-768 770 784 789 797 837 842 848-849 866-867 877 887 893-894 903 914-915 925 929-930 937 944-945 947-949 955 961 967 984
adult lung	GIBCO	ALG001	1 3 14 18 28-29 38 54-56 59 92 110 114-115 130-131 146 149 156 159 164 167 176 184 209 217 234-236 240 255-256 258 263-264 269 271 276 280-281 297 305 308 312 314 322 325 332 336 344 353 361-362 388 401 410 420-421 426-427 431 465 469 474 484 498 500 506 508-509 517 530 532 573 592 596 613 619-620 623 626-628 638 658 679 681 684 689 717 731 741 771 791 799 817 834 845 861-862 864 875-876 901 921 925 928 932 940 947 949 959 962-963

			967
lymph node	Clontech	ALN001	3 10 110 146 160 168 196 209 221 269 278 301 336 348 394 405 411 420 422 459 464 474 485 503 506-507 532 563 582 619 623 630-631 642 669 684 697 713 715 727 747 767 769 789 825 839 842 849 887 896 913 921 925
young liver	GIBCO	ALV001	3 14 16 37-38 41 51 56 60 97 104-105 108 110 117 119 128 130-131 134 139 149 152 169-172 176 184 189-190 200 209 212 216 218 228 232 255 258 263 270-271 275 285-286 292 295 298-299 301 304 314 341 358 365 368 376 400 410-412 431 474 481-482 485 496 500 504-505 517 520-522 524 530 532-533 547 551 563 581 583 610-611 621 624 635 643 691 708 711 715 720 752 755 761 768 796-797 811 818 830 845-847 852 864-865 867-869 896 899 910-911 949 958 965 969 972-973
adult liver	Invitrogen	ALV002	3 37 42 56 60 71 82 104-105 114-115 117-118 125 130-131 134-135 164 169-172 176 179 200 203-204 212 217 223 226 232 237 244 263 274-275 292 301 310-312 314 317 349 354 364 368 372 376 398-399 402 426-427 439 442 451 458 465 474 482 485 490 506 515 525 527 545 547 552 568 571 573-575 582 587 594-595 604-605 608 610 621 630-631 634-635 637 657 664 690 693 699 723 726 745 751 763 767 784 793 811 822 845 848 852 856 861-862 864 892 899 908-909 925 950 958 967 983
adult liver	Clontech	ALV003	60 134 169-171 275
adult ovary	Invitrogen	AOV001	1 3 9-10 12-14 16 18 20 22-25 28-29 33-35 37 39 41-42 46 48-50 55-57 59 63-67 69 71-72 75 77-80 82 88-89 92 101 103-106 108-110 113 115 119-121 123-126 128-133 135 138 142-146 149 151-152 159-161 167-168 172 174 176-177 179 181 184-190 194 198 200 203 208-209 211-212 214 217 219 221 224 226 232-235 240-242 246-247 249 251 254-255 258-259 264 269-271 274 276-277 279-283 285 288 290 293-294 297 301-304 306-308 311 314 319-322 325-326 328-329 331-332 335-338 341-342 344 348 354-358 361-363 365 368 370-372 374 376 379-380 382-383 388 394-396 398-399 401-402 405-406 409-412 416 418-421 423 425-433 438 442-443 449-452 454 462 464 466-467 469-471 474 479

			482-484 488 490 492-496 498 500-504 506-509 511 515-518 520-524 529-530 532-533 537 539-542 545 551 555 558 560-565 569 571 573 577-578 581-583 585-590 592-593 596-597 600-605 608 610-611 613-614 617-628 633-637 639 642-643 646-648 650 652 654 656 658 664 668-670 672 674 679 681 684 688 691 693 697-699 701-702 713 717 721-722 724 729-732 738-744 747-750 752-753 755 759 761 765 767-774 779-780 783-784 789 793 795-797 801 813-818 823-824 828 830-832 834 837 839 841-842 845 848-851 856 859 862 864 866-867 870-871 874-878 881-883 887-889 891 893-894 896-897 901 903 906-911 913 919-922 925 928 930 936 939-940 943-944 946-947 949-950 952-953 955 957-958 962-963 965 967 969 971 973 977 981-982
adult placenta	Invitrogen	APL001	41 56 67 253 301 304 334 380 383 451 474 479 500 577-578 643 648 729 767 856 859 866 873 962-963
placenta	Invitrogen	APL002	3 21 31 38 63-64 78 135 143 168 186-187 212 232 244 263 280-281 334 336 344 348 371 374 394 399 461 490 582 588 602-607 610 620 699 745 769 793 817 822 859 897-898 923 928 931 943 949 969 973
adult spleen	GIBCO	ASP001	1 3 21-22 46 52 54-55 57-58 61-62 72 74 78 82 88 118 121 130-131 137 152 159 168 172 189 203 209 217 223 234-235 252 255 263 269 271 274 282 288 290 301 314 322 335 350 363 394 403 405-406 410-412 415 431 459 464 472-474 482 488 500 506 510 514 517 532 537 542 561-563 589 593 602-603 610 613 619 621 636 642-643 655 658 662 674 676 679 681-682 684 689 691-692 697 699 715 720 723 729 747-748 769-770 782 793 818 830 834 845 856 859 862 877 887 893-894 896 903 906-907 914-915 918 925 928 930 940 946 965 967 977 982
testis	GIBCO	ATS001	6 22 28-29 33-34 41 48 52 62 65 72 97 106 109 118 132-133 145-146 168 172 176 183 185 189-191 195 209 211-212 214 221 223 230 254-255 258 263 269 283 297 312 314 321 342 352 361-362 365 380 383 388 395 401 405-406 412 430-431 441 469-470 474 479 495-496 500 506 520-521 533 543 545 548 560

			563 574 582 589-590 593 608 616-618 620 623-624 638 642-643 697 699 708 711 745 747-748 765 767-768 779 784 789 812-813 834 837 839 848 859 862 868-869 875-877 887 889 893-894 896 928 944 947 953-955 972 981
Genomic DNA from BAC 63I18	Research Genetics (CITB BAC Library)	BAC001	515
Genomic DNA from BAC 393I6	Research Genetics (CITB BAC Library)	BAC002	640
Genomic DNA from BAC 393I6	Research Genetics (CITB BAC Library)	BAC003	640
adult bladder	Invitrogen	BLD001	50 55 66 71 111 143-144 148 160 201 209 223 255-256 280-281 286 305 315 319 340 394 431 442 488 497 505 518 552 588-589 621 636 664 676 715 738-739 769 790 824 837 845 877 887 936 940 948 962-963 967
bone marrow	Clontech	BMD001	3 10-13 16 18 20-21 25 28-29 31-34 41 45 48 52 54-55 57 59 61 65 67 72-73 75 78 80 82 84 99 103 108 110 114-115 118-120 123-124 128 130-133 143-144 148 152 159-161 163 168 172 174 176 178 190 192 198 203 209 211 217-218 221 223-224 227 233-236 244 247 249 252 254 258 260-262 267 269 272 278 280-281 284-285 288 290 294-297 301 304 308 314 317-318 320-321 325 328-330 333-335 349 351-354 358 363 365 367 377 382 388 394-397 400 405 408 410-412 418-421 425-428 431 433 435 442 449-450 453 455 459 464 468-470 474 478-479 481 484 490 496 504 506 508-509 511 519-521 530 532 539 553 558-559 561-563 580 582 586 592 599 608 610 613-614 617-619 623 625-628 635 638 641-643 658 664 672 682 699 711 713 717 731 734 740 742-743 745 761 768-771 774 776-778 784 787 789 813 817-818 822 834 839-840 842 848 862 866 870 876 885-887 891 896-898 900 903 906 913 919 921-922 927-928 939 944 947 950 953 959 961-963 967-968 970 973 977
bone marrow	Clontech	BMD002	3 9-10 15-19 30 33-34 39 45 54 57 63-64 71 82 102 116 119 130-133 148 152 156

			159-160 168 176 182 224 254-255 271-272 282 285 290 297-299 301 305 323 333 340 344 351-355 358 361-362 364 367 370 372 387 394-395 399 403 405 409 411 449-450 459 461 468 474 488-489 524 530 532 580-582 592 602-603 611 617-618 621-622 630-632 642 661 663 694 717 730 734 740 745 752 755 761 767 769-771 775-778 784 787 811 813 818 832 840 842 849 859 878 887 893-894 896-898 903 906 908-909 923 928 944 946-949 953 958-963 965 982
bone marrow	Clontech	BMD004	54
bone marrow	Clontech	BMD007	766 887 928
adult colon	Invitrogen	CLN001	22 37 67 97 117 121 148-149 168 172 190 200 204-205 232 244 263 268 292 301-302 363 377 384 452 455 459 470 530 582 602-603 619 687 723 728 751 761 831 861 887 914-916 934 955 969 984
Mixture of 16 tissues – mRNAs*	Various Vendors*	CTL016	358 740 760
Mixture of 16 tissues - mRNAs*	Various Vendors*	CTL021	468 527 928
adult cervix	BioChain	CVX001	1 3 10 14 22 28-30 37 41 47-48 51-52 54-57 71 82 89-90 92 106 108 110-111 117-118 121 129-131 135 141 143-146 160-161 164 168 172 177 189-190 193 195 200 204 209 211-212 217 226 229-230 232 234-235 240-242 246 254 260-263 268-270 274 277 282 285 292 295 297 305-308 314-316 319 328 343-344 348 354 358 363 368 380 382-384 389 394 396 399 401 405-407 410 416 418-421 428 430-431 437 442 453-454 459 464 469 471-473 476 480 484 492-495 500 504 506-509 516-517 526 530 532 545 550-551 563-565 569 577-578 585-586 590 608 611 613 619 621 623 628 630-631 634-637 641 643 648 656-658 664-665 674 679 682 689-690 693 700 703 708 713 721-722 724 728 732 742-743 747 750 752 755 757 761 763 767-769

* The 16 tissue-mRNAs and their vendor source, are as follows: 1) Normal adult brain mRNA (Invitrogen), 2) normal adult kidney mRNA (Invitrogen), 3) normal adult liver mRNA (Invitrogen), 4) normal fetal brain mRNA (Invitrogen), 5) normal fetal kidney mRNA (Invitrogen), 6) normal fetal liver mRNA (Invitrogen), 7) normal fetal skin mRNA (Invitrogen), 8) human adrenal gland mRNA (Clontech), 9) human bone marrow mRNA (Clontech), 10) human leukemia lymphablastic mRNA (Clontech), 11) human thymus mRNA (Clontech), 12) human lymph node mRNA (Clontech), 13) human spinal cord mRNA (Clontech), 14) human thyroid mRNA (Clontech), 15) human esophagus mRNA (BioChain), 16) human conceptional umbilical cord mRNA (BioChain).

			779-780 784 788 810-811 813-815 822 834 836-837 839 848 861 866-867 871 874 877 887 891-894 897-898 901 913 916 919 921-922 925 946-947 953 958-959 967 969 973
diaphragm	BioChain	DIA002	3 39 184 203 431 563 848 967
endothelial cells	Strategene	EDT001	3 6 8-10 14 19-24 28-29 33-34 37 39 41 46 48 52 55-58 62-65 67 69 71-72 75 78 80 82-83 87 101-102 108-109 114-115 117 123-124 128 130-133 135 138 143 145-146 149 156 159-160 167-168 172 174 176-177 179 181 184-187 189-190 194-195 200 203 208-209 212 216-217 219 223-224 226-227 229 234-235 244 248-249 254-256 258 263-264 267 269 271 274 276-282 285 290-291 294 297 301-304 308 311 313-314 316-317 320-321 323 325-326 328-329 331-332 334-337 339-341 344 348-349 352 354-355 358 361-363 365 367 371-372 375 379-380 383 389 394-395 398-403 405-406 409-412 425-428 437 442-443 448 454 464 466-467 474 479 481 490 492-498 500 503 506-509 511 517 520-521 523-524 530 532 537 540-542 558 561-563 565 569-570 573 581-583 586 588-589 596 602-608 610-611 613 617-622 625 628 630-631 633-637 642-643 646 648 650 652 659 661-662 682 688 690-693 696 698-699 708 712 715 717 720-722 724 727 729 740 745 748-750 752 761 765 767-770 772-773 779 784 789 792-794 796 802-803 811 817-818 821 824 827-828 830 834-835 837 842 845 848 859 861-862 864 866-867 870 876 885 887 891 893-894 897-898 900 903 906-907 913 916 921 925 939 947 950 953 955 957-958 962-963 967 973 978 984
Genomic clones from the short arm of chromosome 8	Genomic DNA from Genetic Research	EPM001	324 515 640
esophagus	BioChain	ESO002	97 103 128 371 474
fetal brain	Clontech	FBR001	67 129 156 159 232 267 433 446 503 845 952
fetal brain	Clontech	FBR004	28-29 185 213 277 350 384 432 485 501 549 651 747 754 761 780 787 848 870 887 906 958
fetal brain	Clontech	FBR006	10-11 14 21 30 32 47 49 56 65 69 72 77-78 82 84 97 101 115 118 121 125 128 130-131 138 142 148 152 159-160 179 185 188 194 197 203 210 212 214 219

			222 227-229 243-246 249 252 256 264 270 273 282 285 290-291 293 301-303 305-306 312 321-322 325 327 339-340 344 346 350 354-357 363 367-371 374 388 391 394-395 399 402 405-406 410 414 420 426-427 436-437 442 444 454 456-457 460 462 464 470 480 485 492-494 507 510 516 524 528 530-532 539-542 549 553-554 561-562 580-582 588-589 602-608 611 615 617-619 621-622 624 632 636 641-642 646-647 651-653 661-662 666-669 672 677 691 715-716 730 735 740 752 754 761 767-770 772-775 780-781 799-801 808 818 822-823 835 843 845 856 859 864 867 876 880 885 887 890 893-894 896 913 918 926 942 946-947 951 957-959 962-963 970-971
fetal brain	Clontech	FBRs03	130-131 312 517 637 691 738-739
fetal brain	Invitrogen	FBT002	3 22 28-31 47 57 63-64 72 75 77-78 86 94-95 97-98 126-127 135 140 143 156 159-160 167-168 177 185 190 196 201 203-204 214 217 230 254-255 258 267 273-274 277 279 282-283 292 301-302 305 312 314 323 329 346 348 367 374 382 394 399 401 403 412 415 420 432 437 474 482 485 495 507 513 517 527 529-530 539-542 548 552 579 587-588 600 604-605 612 617-618 621-622 624 634 642-643 647-648 650 679 689 693 699 712 715 742-743 745 748-749 753 768-769 793 797 829-831 834 845 848 856 859 893-894 908-909 913 916 931 933 940 950 967 969
fetal heart	Invitrogen	FHR001	19 57 130-131 394 431 642 769 844
fetal kidney	Clontech	FKD001	3 31 33-34 38 48 54 72 160 208-209 211 223 264 269 277 283 290 313 325 341 348 358 396 418-420 474 484 506 508-509 517 520-521 532 547 553 558 567 569 587 596 608 610 613 619 622 626-627 642 679 734 745 818 843 887 896 903 916 969 971
fetal kidney	Clontech	FKD002	19 474 726 903
fetal kidney	Invitrogen	FKD007	3 118 186-187 230 244 271 432 887 969
fetal lung	Clontech	FLG001	69 132-133 156 168 208-209 217 267 269 274-275 286 354 394 396 406 462 483-484 608 619 751 769 771 834 914-915 925
fetal lung	Invitrogen	FLG003	3 8 28-29 32 39 50 66 82 88 92 168 186-187 200 204 212 226 229 246 274 309 327 332 368 374 382 394 398 426-427 431-432 442 485 536 555-557 587 604-

			605 621 624 636 642-643 661 677-678 724 753 769 848 859 864 877-878 896 902 904 914-915 958
fetal lung	Clontech	FLG004	130-131 394 664 769 942
fetal liver-spleen	Columbia University	FLS001	3 8-10 12-13 16-17 19-25 27-29 33-35 37-38 41 45-46 48 52 55-58 60-67 69 71-74 77-78 80 82 84 87-90 104-106 108-109 112-121 123-125 128-134 138 141 143-146 149 151 156 159 163-164 167-172 174 176-179 181 184 186-188 190 194-200-201 203 208-209 211-212 216-217 219 224-227 229-230 232 234-235 237 241 243-244 246-248 254-255 258 260-263 267 269-270 273-282 284-285 288-290 292-295 297-299 301-306 308 311-318 320-323 326 328 332 335 341-344 348 352 354-359 361-365 367-368 371-374 376-380 382-383 388-389 394-396 398-399 401-411 413-414 416 418-421 425 428-430 432-433 437 439 442-444 449-450 452 456-457 461-470 472-474 478-479 481-482 484-485 487 490-494 497-499 504-507 511 514-515 517-521 523-524 526 529 532 537 540-541 547 555 558-559 563 575 577-578 580-596 598-599 601-603 606-608 610-613 617-624 626-628 630-631 634-636 639 642-643 647-648 654-656 663-665 672 674-675 679 681 684 686 688 691 693-699 711 713 715 717 719-726 729 732-733 738-740 745 748-749 751-753 757 759 761 767-770 776-778 780 784 787 792-794 799 804 809 811 813 817-819 822-825 830-831 834 837 840 842 845-848 852 856 859 861-862 865 867-869 871 874-878 887-888 891 893-894 896-900 903 905-911 913 916 918 923 928 930-931 936 939 942 944 946-950 952 958-959 961-963 965 967 969-970 972-973 976-977 981-983
fetal liver-spleen	Columbia University	FLS002	3 8-13 15-17 19-20 22 25 28-29 33-35 37 41 45-46 52 54-56 60-61 63-64 66-70 73-74 78 80 82 92 99 104-106 108-109 112 115-116 118 120-121 123-125 128 132-135 139 141 143-144 146 149 152 156 159-161 167 169-172 174 176-177 179 181 185 188 190 194 196-197 200 204 212 214 216-218 223-224 226-230 232-235 237 246-247 252 254-255 258-263 267 270-277 284-286 288 292 294-295 297-299 301 303-305 308 310 314 318 320 323 328 330-332 335-337 340 342-

			344 352 354-355 358 361-365 367-368 371 373-374 376-377 382 388 394-396 398-399 401 405-406 409-411 413 418-421 429 431 439-440 442-444 451-452 457 462-463 466-468 470 474 477-479 481 483-484 487-488 491 495 499 504 508-509 516 519-521 524 526-528 530 532 537 540-541 543 545-547 550-551 553 555 560 564 568 574-575 577-578 580-592 596-597 600 602-603 608 610-611 613-614 617-618 621-622 628 630-631 634 637 639 642 644 647 654 658-659 665-667 669-675 679 681 684-685 688-690 693 695 697 708 711 713 715 717-719 723-727 729 731-734 738-739 741 745-746 749-750 753 759 761 766-767 769-770 776-779 782 784 791-792 794 805 808 817-818 822 824-825 830 834 837 842 845-849 852 856 859 864-865 867 874-878 888 891-892 896-900 903 905-906 908-909 913 916 918 921 923 925 932 936 939-940 942 944 946-947 949-950 953 955-956 958-959 961-963 965 968-970 973 977-978 981
fetal liver-spleen	Columbia University	FLS003	19 60 78 224 273 275 370 373-374 401 602-603 639 643 730 732 738-739 748 752 770 782 928 930 947 949
fetal liver	Invitrogen	FLV001	37 55 60 69 72-73 97 104-105 108 113-114 116-118 121 135 143 152 167-168 186-187 195 200-201 209 217 223 240 244 253 255 275 284 301 311 314 317 336 342 348-349 358 371 374 382 394 402 411-412 418-419 428 430 442 453 517 568-569 580 582 584 587 589 601-603 606-608 617-618 624 634 639 642-644 646 664-665 669 679 715 717 720 726 745 748 751 769-770 782 791 794 797 824 830-831 845-847 852 859 870 899 913-916 925 928 948 956 958 969 976 982
fetal liver	Clontech	FLV002	72 418-419 632
fetal liver	Clontech	FLV004	3 160 169-171 355 367 374 376 547 617-618 621 646 717 741 771 836 878 976
fetal muscle	Invitrogen	FMS001	15 27 32 37 67 72 83 99 112 121 138 167 174 177 186-187 190 203-204 211 215 230 252 259 312 374 403 406 409 457 461 485 505 517 528 530 540-541 544 549 554 558 579-580 583 602-603 608 639 642-643 654 664 699 715 730 737 751 772-773 788 802-803 810 848 856 859 864 868-869 887 893-894 905-906 910-911 923 948 967

fetal muscle	Invitrogen	FMS002	15 99 130-131 223 361-362 431 474 505 581 639 643 666-667 784 790 808 810-811 874 880 887 903 946 950 958 962-963 973
fetal skin	Invitrogen	FSK001	3 6 20-22 32-34 41-45 47 49-52 55 63-64 66 69 77 80 88 91 98 101 111-112 115 126 130-131 135 142 144 146 160 163 167 176 188-190 196 201 204 208 213 215 217-218 229 232 244 246 248 255 263 265-269 274 279-281 283 285 288 292 294 297 301 303 308 314 321 341-342 344 348 354-355 358 361-362 366 369 371-372 374 381-382 384 386 394 401 403 405 413 415 428 431 437 440 460 466-467 472-473 477 481 483 495 499 504 517 522 532 536-537 539-541 545 556-558 569 574 576-578 580 584-585 587-589 592-593 602-603 606-608 612 617-618 621 624 634 637 639 642-643 647 664 673-674 676 680-681 689 699 705-707 709-715 724 728-730 738-740 745 748 752 765 768-769 772-773 793 797 817 823 830 834 842 848 859 861 864 870 874 883 887-888 893-894 901 904 908-909 913-916 923 925 947 950 958 962-964 967 975
fetal skin	Invitrogen	FSK002	3 130-131 146 194 306 354 367 400 405 474 489 520-521 547 558 561-562 585 596 730 740 748 755 767 771 810 840 893-894 946 959
fetal spleen	BioChain	FSP001	276 563 842
umbilical cord	BioChain	FUC001	3 20 33-34 39 48 50 52 55-57 65 67 69 72 77 79 82 92 109 112-113 121 132-133 138-143 156 167-168 172 174 179 184-185 190 194-196 200 202-203 208-209 229-230 244 269-271 278 284-285 290 297-299 303 305 308 320 331-332 336 338 342-343 363 367 372 374 379-380 383-384 392-394 397 399 402 405-406 410 425-427 429-430 449-450 474 476 484 497 499 501 504-505 510 515 517 532-533 539 549 551 558 563 569 574 577-578 581 586-587 597 602-603 608 610 617-619 621 626-627 634-637 639 642-643 658 663-664 674 690-691 693-694 699 713 715-717 720 724 726 729 738-739 746-747 749 759 761 765 768-769 774-775 793 797 807 818 822 837 848-849 856 862 868-869 874 885 887 892-894 903 906-907 916-917 919-920 928 936 939 944 946-947 962-963 967 969

fetal brain	GIBCO	HFB001	3 9-10 12-14 16 21 25 28-30 32-34 37-39 41 47-48 52-53 56 65 67 69 71-72 75 80 84 92 97 103 106 110 114 117-119 123-124 127 129 132-133 135 138 141-142 144-146 148-149 152 156 159-160 168 172 174 176 179 181 184-185 190 198 208-209 212 214 219 221 223-224 229-230 233-236 240 244 247 251 253-255 258-259 270 273 276-277 285 297 304-305 308 312 314 322-323 325 328 332-333 335-337 339-340 342-344 346 352 354 358 363 365 370-372 374 382 394-396 398 401 403 405-406 409-412 414 416 425-427 431-432 437 442 445 453 456 462 466-467 469-470 472-474 479 483 488 490 492-497 500-501 504 506-510 520-521 524 530 537 539 545 549 552 558 560-562 564 569 579 582-583 586-587 596 602-608 610-612 614 617-624 626-628 630-631 633 635 638 641 643 647-648 656 658 661 676 679 688-689 693 696-697 711-712 715 724 726 731 735 745 747-749 752 754 761 765 767-770 774 779-781 784-786 789 799-800 802-803 813 818-819 823-824 831 834-835 837 839 845 848 859 864 866-867 871 874-875 881 887 891 893-894 896-897 900 906-907 910-911 918 921-922 925 927-928 930 943-944 946-947 950 953 962-963 965 969 972-973 977
macrophage	Invitrogen	HMP001	86 168 186-187 297 537 608 681 761 845 877
infant brain	Columbia University	IB2002	2-3 9-10 12-14 16 21 25 27-30 32 37-38 46-47 49 55-56 58 65 69 71-72 78-79 82 84-86 91-92 98-99 106 109-110 113-115 118 127-128 130-133 135 138 142 144 151 156 168 173-176 180-181 185-188 192 194 196-201 203 208 210-212 214 217-218 224 229-231 233 236 238 240-241 244 246 251-256 259 263 270-271 277-279 284-285 287 293-294 296 301-302 308 312-314 317 322-323 327 330 333 339 342 345-346 351 354 358 361-362 365-366 368 370-371 373-374 382 388 394-396 402 405-406 411-412 415-416 420 424-425 428 431 436-437 440-441 444-445 453 456 460 465 474 479 482-483 488 495-496 498 501 503-504 506-510 515-517 520-521 524-525 529 531-532 534-535 537 539-542 544-545 549 561-562 569 574 577-578 580-583 586-587 589 592 596 600-608 610 612-

			613 616-618 620 622 624 629-632 634-635 637 641 643-644 650-651 653 661 663-664 676-677 689 693 695-698 708 711 720-722 724 730 732 735 740 745-748 754 765-766 768-769 779-781 785-786 789 791 796 798 800-803 807 811-813 818-819 822-824 830-831 834-835 837 839 842-843 845 854 856 858 864 867-869 875-877 879 881 887 892-894 896 903 907-911 913 916 919-920 925 930-932 936 939 943 946-947 953 958 970-973 977-978 982 984
infant brain	Columbia University	IB2003	3 12-13 21 27-29 32 39 49 69 72 82 91 113 116 126 128 132-133 142 144 156 176-177 184-185 188 194 208 212 223-224 228 230 244 255 259 267 270 273 276 293-294 312 320 326-327 337 342 346 354-355 358 361-363 382 388 390 394 396 399 402 420 425 431 442 462 474 482 484 488 495-496 510 520-522 524 529 540-541 549 563 582 586 588-589 596 600-603 606-607 612 617-618 620-621 632 647 650 679 720-722 724 735-736 746 751 754 769 785-786 793 800 807 811-813 818-819 822 824 831 834 838-840 843 856 864 892 896 907 919-920 925 930-931 936 947 950 957 973 982
infant brain	Columbia University	IBM002	16 47 82 84 201 263 302 376 394 421 440 488 537 592 606-607 635 740 769 887 892 906 921 926 971
infant brain	Columbia University	IBS001	84 86 180 185 198 201 203 230 279 312 326 346 354 366 388 488 542 581 588 620 647 664 732 740 785-786 801 807 822 827 910-911 925 931
lung, fibroblast	Strategene	LFB001	3 11 25 49 65 75 114 141 156 160 172 190 198 209 217 224 229 234-235 267 269 274 277 282 284 303 308 312 320 334 336 352 372 396 398 412 414 437 453 464 470 481 492-494 508-509 532 539 581 584 617-619 621 628 633 643 688 691 745 752 761 768 794 822 837 848 876 887 953 967 973
lung tumor	Invitrogen	LGT002	1 3 9-10 12-13 20 31 38 41 46 48 51-52 56 58 63-64 72 74-75 78 82 88 101 106-107 110 114-115 117-118 120-121 123-124 128-133 135 143-146 149 151 156 159-161 163-164 167-168 172 176 178-179 184-185 189-191 194-196 200 203 209 212 216-217 226 228-229 232 234-236 241 246 248 256 258-259 263-264 269-271 274 282-283 285-286 290 292

			294 297 301 308-309 311 314 317 321 326 328-329 331 333-334 341 348 352 354-355 363 365 371 380 382-383 388 394-395 398-402 405-406 410-411 413 416 418-419 426-427 439 442 452-453 458-459 461-462 464-465 470-471 474 478 483-484 490 495-496 499 510 522 524 528 536-537 540-541 543 548 556-558 560-565 571-573 580 582 587-588 592 597 602-605 608 610 612-613 617-622 625-629 633-634 636 642-644 648 661 664 669 679 688-689 691 693 699-700 708 717 723-724 730 733-734 738-740 745 747 749 752-753 761 767-768 770 779 782 784-786 789 793-794 797 817-818 820 823-824 834 837 842 845 848 855 857 859 862 864 866 870 875-877 887 892 896 900-901 907-909 914-915 919-920 923-925 939 943 947 949 953 958 962-963 965 968 970 972-973 977
lymphocytes	ATCC	LPC001	3 9-11 32 47 50 56 71 75 88 97 99 102 121 125 128-129 135 138 141 149 163 167-168 212-213 217 233 255 290 294 301 305 311 314 342 372 377 388 398-399 410 437 442 453 470 474 481 495 500 506 510 529 532 537 542 558 571 579 604-605 610 620 628 637 643 658 666-667 676 679 697 708 713 728 730 734 749 765 768 796 807 818 822 834 839 848 859 875 885 887 896 903 906 914-915 928 947 973 981-982
leukocyte	GIBCO	LUC001	1 3 9 11 18-19 21 23-25 27 31-34 39 41-42 46-48 52 54-58 62-69 71-72 74-75 78-80 82 89-90 93 99 110 115-121 123-124 128-133 135 138 141 143-146 149 152 156 159-161 163 167-168 176 179 181 186-187 189-190 194 198 200 203-204 209 211-212 218-219 226 232-236 240 244 247 251 253-255 258-259 263-264 269 271 274 278-279 282-283 285 288-290 294-295 297 301-306 311 313-314 317 320-321 325 328 330-331 335 337 342 344 348 350-351 353-354 358-359 361-365 368 371-372 375 388-389 394-395 397-401 403 405 407 409-412 421 425-427 432 437 442 448-450 452 457 460-461 468-471 474 476 479-482 484 492-494 496-498 500 506-510 516-517 520-521 524 529-530 532 537 540-544 551 553-554 558 560-565 569 577-578 580-583 586-587 589 592 596-597 602-

			603 606-608 610-624 626-628 630-631 634-635 641-643 654 657-658 661 663-665 669 672 677 679 684-689 691 696-697 699 708 711 713 715 717 721-724 728 730 738-740 747-749 755 761 765 767-769 771 774-779 782 784 789 791-792 794-795 797 807-808 811-815 817-818 822 824 828 830 832 834 839-840 842 845 848 856 859 862 864 867 871 875-877 887 891 893-894 896-898 903 906-911 913-916 921 923 925 927-928 930 932 935-936 939 943-944 947 949-950 953 958-959 961-963 965 967 972-973 982
leukocyte	Clontech	LUC003	1 41 82 106 119 123-124 160 177 184 201 212 221 228 271 279 285 295 321 325 372 394 411-412 443 468-470 530 532 537 551 569 580-581 613 619 623 626-627 642 655 697 761 767 769 775 789 809 867 887 923 928 950
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	3 25 55-56 67 71 78 109 121 129 146 167 172-173 176 200 209 212 258-259 263 278 297 301 306 312 335 338 340 352 361-362 367 388 395 402 410 418-419 429 437 454 464-465 481 496 500 503 507 524 532 539 560-562 581-582 587 589 599 612-613 617-621 623 643 657 663-664 672 715 724 748 752 761 767-768 770 785-786 789 835 848 877 887 896 916 919-920 947 967 978-980
mammary gland	Invitrogen	MMG001	1 14 19 21 28-29 31-37 47 49-51 55 57 63-67 69 71-72 75-78 92 108-109 111 116 121 123-124 126 128 130-133 135 143-144 148-150 156 159 164 168 172 177-179 184 186-187 190 194 200-204 209 212 217 226 230 232-236 241 244 246-247 252 255 258-259 263 268 270 275 279-283 285 290 292-293 301 304-305 311 313-314 317 320 322-323 326-327 330 332 338 342-344 348-349 354 360 363 367 371 374 380 382-383 385 388 394-395 398 401-403 407 409 411-412 418-420 426-427 430 435 437 442 449-453 459 461 465-468 470 474 477-478 480 483 485 488 498 500 503-504 507 515 519 522 524 529-532 538-541 544 547 555 560 563 565 569 573-574 579-580 582 584 587-589 593 597 601-610 612-613 615-618 620-622 624 634 636-637 639 642-644 646-647 650 657 663-664 674 676 679 688-689 691 693 696 701-703 713 715 717 728 730 732 738-

			739 741-743 745 749 751 753 763 767 769 772-773 785-786 793 796-797 812 821-824 830-833 837 848 856 859 861 864 868-870 876-877 887 891 893-894 898 903-904 907-911 913-918 921 923 925-926 930-931 936 942 949-950 958 961 966-967 969 972-973
induced neuron cells	Strategene	NTD001	9 65 82 92 106 113 142 146 156 172 176 191 208 221 258 277 328 333 346 361-362 371-372 375 388 410 414 418-419 440 471 484 495 516 524 529-530 592 610 628 642 650 745 748 752 761 793 818 848 851 897
retinoid acid induced neuron cells	Strategene	NTR001	19 87 184 305 385 440 474 626-627 643 748 799 834 977
neuronal cells	Strategene	NTU001	19 33-34 42 70 82 87 109 115 126 146 172 185 188 194 212 255 269 274 283 312 317 329 340 361-362 367 379 394 399 401 410 420 426-427 474 479 507 530 579 582-583 610 617-618 636 643 658 732 740 765 769 784 791 793 799 802-803 818 842 851 864 897 907 932
pituitary gland	Clontech	PIT004	3 19 123-124 194 255 354 358 373-374 377 426-427 462 492-494 635 785-786 793 893-894
placenta	Clontech	PLA003	138 176 574 896 972
prostate	Clontech	PRT001	3 9 16 57 65 75 83 108 130-134 138 141 146 149-150 159 182 186-187 190 203 209 234-235 276 283 322 413 415 442 449-450 453 480 484 490 499-500 503 505-506 523 537 543 564 583 602-603 611 619 623 643 650 697 711 729 761 765 770 776-778 784 789 819 822 831 839 862 866 887 904 907 921 935 962-963 967 973
rectum	Invitrogen	REC001	19 30 33-34 66 108-109 123-124 126 129-131 143 149 151 156 164 190 201 240 247 250 263 268 274 279 287 295 298-299 310 314 332 341 354 384 394 401 420 425 442 446 459 483 485 520-521 532 545 559 580-581 584 592 602-607 610 612 615 619 634 637 646 655 664 683-684 741 769 793 822 870 908-911 914-916 934 937-938 942 967 973 982
salivary gland	Clontech	SAL001	16 68 74 84 121 123-124 156 172 190 203 209 232 248 254 269 292 294 363 377 395 398 400 402 405-406 410 430 442 459 462 474 483 485 563-564 579 587-588 599 602-603 643 658 699 728 730 737 741 748 794 822 867 876 897 903 981

salivary gland	Clontech	SALs03	217 254 270 388 610
skin fibroblast	ATCC	SFB001	517 949
skin fibroblast	ATCC	SFB002	269 688
skin fibroblast	ATCC	SFB003	3 203 897 907
small intestine	Clontech	SIN001	3-4 47 57 68-69 92 99 125-126 130-131 135 149 151-152 156 159 185 204 241 246 291-292 318-319 338 343 348 363 373 375 382 388-389 392-394 397 400 437 466-467 471 484 500 517 520-521 525 547 560 580-581 588 599 602-603 612 624 643 711 731 733-734 757 761 769 774-775 794 824 864 904 906 910-911 913 948 953 959 976 984
skeletal muscle	Clontech	SKM001	15 75 135 146 172 190 218 267 282 308 410 426-427 474 505 588 620 623 658 692 713 737 779 790 862 874 878 887 952 962-963
skeletal muscle	Clontech	SKMs04	215
spinal cord	Clontech	SPC001	14 20-21 25 28-29 31 39 46 48 59 78 83-84 91-92 103 112-113 135 160 168 172 176 188 190 205 209 229 232 258 285 301 308 312-314 321 323 329 346 374 377 380 383 388 394 398 406 409-410 431 449-450 453 455 466-467 470-471 484-486 488 495 497 500 503 508-509 524 537 539 558 581 586 604-605 611 619 623 630-631 633 656 663 711 715 729 736 740-741 761 767 769 776-778 780 818 822 831 835-836 840 843 859 861 871 875 887-888 897 906-907 913 919-920 928 931 953 958
adult spleen	Clontech	SPLc01	3 6 12-13 66 130-131 178 365 403 431 461 558 610 715 797 809 876 947 967
stomach	Clontech	STO001	35 114 130-131 144 155 176 189 206-207 249 260-262 336 382 398 425 431 453 461 483 496 500 527 530 580 642 657 663 669 748 765 768 802-803 839 891 942 981
thalamus	Clontech	THA002	30-32 48 66 109 127 130-131 135 142 145 156-158 168 172 174 185 199 224-225 233 246 277 282 286 293 322 332 334 346 374 384 400 402 420 424 435-437 446 466-467 485 503 506 527 542 549 572 612 615 622 624 633 643-644 658 676 736 790 794 824 831 835 896 907 950 969
thymus	Clonetech	THM001	10 16 20 28-29 32 37 41 52 57 66-67 74-75 110 118 121 129-131 141 151 159-160 208 211 218 247 269 289 295 297 320 325 354 358 365 367 372 378 388-389 395 398 411-412 420 423 435 452 500 508-509 517 524 532 537 551 558 560

			569 577-578 582 586 598 608 611 622 643 684 715 721-723 728 740 766 772-773 795 834 837 849 864 885 900 921 946 948 958 962-963 965 972-973 982
thymus	Clontech	THMc02	1 3 9-11 16 21 27 32-34 38-39 51 55-57 66 72 74 77-78 80 82 89-90 101 112 115 118-119 121 123-124 126 138 144 152 159 168 174 176 178 186-188 197 200 208 212-214 217 225 233 243-244 246 254 256-262 279 282 285 288-289 296-297 313-314 322 334 343 354-355 358-359 363-364 367-368 372-373 382 387-389 395 400 402 411 414 426-427 437 440 442 449-450 454 457 462 464 469 474 479 481 485 490-491 506 508-509 511 517 522 526 528 532 542 551 554 561-562 564 566-570 580-582 585 589 597 599-600 602-608 611 613-614 619-621 625 628 630-631 644 646 655 669 672 677 684 686-693 697 713 717 720 728 740 746 749 760-762 767 771 775 794 797 804 808 811 816 818-819 837 840 859 880 883 887-888 896-897 903 908-911 913 916 924 936 947-948 950 962-963 965 967 970
thyroid gland	Clontech	THR001	3 8-9 14-15 19-22 28-29 39 41 55-56 66 69 71-72 78-79 97 104-105 109 113 115 119 121 123-124 130-133 135 138 143-144 146 148 151-152 156 159-163 165 168 172 174 177 183-184 196 199-200 203 209 211 215-218 228-229 232-236 244 254-255 258 273 282 290 292 294 297 303-306 308 311 317-318 322-323 325-326 334-335 340 342 348 354 358 373 377 381-382 387 394 398 401-402 405-406 409-412 416 422 425-427 429-431 440 449-453 462 466-468 474 478-479 481-484 490 492-496 500-501 505-506 517-518 522-525 532 537 540-541 545 551 558 560 563-564 580 583 587-589 593 597 599 606-607 610 617-621 625-628 633 635 641-643 658-659 664-669 674 682 686 688-691 696 699 715 724 730 740 742-743 747 750 752 759 761 765-766 768-769 779 789 796 802-803 813 818-819 822 831 837 843 845 848-849 862 864 868-869 871 874 876-877 887 893-894 896-897 907-909 912 919-921 923 925 928 936 940-942 944 946-947 950 953 955 958-959 962-963 967 969 973 981
trachea	Clontech	TRC001	33-34 55-56 69 74 163 172 190 209 212

			267 270 297 305 314 352 413 426-427 466-467 500 502 504 580 586 610 613 633 642 688 691 711 724 738-739 774 782 816 820 839 848 862 868-869 914-915 928 968
uterus	Clontech	UTR001	4 9 18 37 63-64 74 108 114-115 130-131 160 166 179 184 190 209 233 249 269 285 301 314 327 337 348 384 394 399-400 403 406 411 425 431 434 437 440 462 474 485 490 508-509 526 532 579 617-619 636 642-643 672 761 769 793 837 849 864 887 903 906 928 934 947 967

TABLE 2

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
1	L06175	Homo sapiens	occurs in MHC class I region; ORF	308	98
2	Y70775	Homo sapiens	Follistatin-related protein zfsta.	3094	98
3	X15187	Homo sapiens	precursor polypeptide (AA -21 to 782)	4112	100
4	AF110640	Homo sapiens	orphan seven-transmembrane receptor	344	100
5	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	158	72
6	W85607	Homo sapiens	Secreted protein clone da228_6.	1477	100
7	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	884	88
8	Y15227	Homo sapiens	Leu1	391	100
9	Y28817	Homo sapiens	pt326_4 secreted protein.	3338	100
10	X92106	Homo sapiens	bleomycin hydrolase	2445	100
11	Y15228	Homo sapiens	Leu2	445	100
12	U27838	Mus musculus	glycosyl-phosphatidyl-inositol-anchored protein homolog	432	34
13	U27838	Mus musculus	glycosyl-phosphatidyl-inositol-anchored protein homolog	320	27
14	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	2323	99
15	U96781	Homo sapiens	Ca2+ ATPase of fast-twitch skeletal muscle sacroplasmic reticulum, adult isoform	5145	100
16	M16653	Homo sapiens	pancreatic elastase IIB zymogen	1435	99
17	Y13398	Homo sapiens	Amino acid sequence of protein PRO346.	1749	99
18	Y02283	Homo sapiens	Secreted protein clone br342_11 polypeptide sequence.	1399	99
19	Y53030	Homo sapiens	Human secreted protein clone d24_1 protein sequence SEQ ID NO:66.	1371	100
20	AL031320	Homo sapiens	dJ20N2.5 (novel protein similar to fucosidase, alpha-L-1, tissue (EC 3.2.1.51, alpha-l-fucosidase fucohydrolase))	2597	99
21	B01384	Homo sapiens	Neuron-associated protein.	1876	100
22	Y68778	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-10.	2470	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
23	Y55935	Homo sapiens	Human KHS2 protein.	4781	99
24	Y55935	Homo sapiens	Human KHS2 protein.	2807	100
25	AC024792	Caenorhabditis elegans	contains similarity to TR:O95029	463	31
26	Y07972	787	Human secreted protein fragment	1540	100
27	X97630	Homo sapiens	serine/threonine protein kinase	3781	98
28	AF150755	Mus musculus	microtubule-actin crosslinking factor	3514	68
29	AF150755	Mus musculus	microtubule-actin crosslinking factor	3725	70
30	Z38011	Mus musculus	DMR-N9	2988	86
31	AJ000522	Homo sapiens	axonemal dynein heavy chain	6058	99
32	AF037256	Mus musculus	ES2 protein	2260	91
33	S62140	Homo sapiens	TLS=nuclear RNA-binding protein	2917	100
34	S62140	Homo sapiens	TLS=nuclear RNA-binding protein	2890	98
36	AB038237	Homo sapiens	G protein-coupled receptor C5L2	1767	100
37	D79994	Homo sapiens	similar to ankyrin of Chromatium vinosum.	6089	99
38	X63380	Homo sapiens	serum response factor-related protein	1966	99
39	AL022072	Schizosacchar omyces pombe	lipoic acid synthetase	1067	61
40	J03930	Homo sapiens	alkaline phosphatase	2751	100
41	AF132968	Homo sapiens	CGI-34 protein	1088	98
42	AL117637	Homo sapiens	hypothetical protein	2208	100
43	AL021393	Homo sapiens	bK747E2.1 (novel protein)	1526	100
44	X68011	Homo sapiens	ZNF81	1886	100
45	AC002464	Homo sapiens	organic cation transporter; 50% similarity to JC4884 (PID:g2143892)	2423	100
46	W78245	Homo sapiens	Fragment of human secreted protein encoded by gene 19.	1949	100
47	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3604	100
48	AF097330	Homo sapiens	H1 chloride channel; p64H1; CLIC4	1305	99
50	U09413	Homo sapiens	zinc finger protein ZNF135	1361	57
51	AF061812	Homo sapiens	keratin 16	2374	100
52	W63681	Homo sapiens	Human secreted protein 1.	1326	99
53	AB035303	Homo sapiens	cadherin-10	4094	100
54	A12022	synthetic construct	MRP-8	485	100
55	AL121897	Homo sapiens	bA392M18.3 (KIAA0180)	1867	100
56	Y73330	Homo sapiens	HTRM clone 397663 protein sequence.	818	96
57	AF151018	Homo sapiens	HSPC184	955	100
58	AF125042	Homo sapiens	bisphosphate 3'-nucleotidase	1586	100
59	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
60	X04494	Homo sapiens	precursor polypeptide	1903	100
61	AF208865	Homo sapiens	EDRF	528	100
62	D15057	Homo sapiens	DAD-1	567	100
63	AF260665	Homo sapiens	histone acetyltransferase	1510	100
64	AF260665	Homo sapiens	histone acetyltransferase	1429	96
65	AJ277145	Homo sapiens	ras-related small GTPase RAB18	1073	100
66	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein sequence SEQ ID NO:106.	348	100
67	Y82744	Homo sapiens	DNA replication and repair associated protein (DRASP).	1028	100
68	Y44486	Homo sapiens	Human GPRW receptor polypeptide.	1721	100
69	AL031228	Homo sapiens	dJ1033B10.2 (WD40 protein BING4 (similar to S. cerevisiae YER082C, M. sexta MNG10 and C. elegans F28D1.1)	3196	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
70	AJ276316	Homo sapiens	zinc finger protein 304	1751	52
71	Y18314	Homo sapiens	paraplegin-like protein	4146	99
72	AF157028	Homo sapiens	protein phosphatase methylesterase-1	2017	100
74	Y71082	Homo sapiens	Human B-aggressive lymphoma (BAL) protein.	1765	99
75	AF225420	Homo sapiens	AD025	734	100
76	X95235	Homo sapiens	transcription factor AP2	217	100
77	AF108420	Takifugu rubripes	1-aminocyclopropane-carboxilate synthase	733	56
78	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
79	AL117635	Homo sapiens	hypothetical protein	922	99
81	Z85986	Homo sapiens	dJ108K11.3 (similar to yeast suppressor protein SRP40)	865	77
82	AF183414	Homo sapiens	hemin-sensitive initiation factor 2a kinase	3231	99
83	G01143	Homo sapiens	Human secreted protein, SEQ ID NO: 5224.	495	98
84	U03985	Homo sapiens	N-ethylmaleimide-sensitive factor	3744	99
85	Y17791	Homo sapiens	VAX2 protein	1496	100
87	AF263538	Homo sapiens	growth differentiation factor 3	1944	99
88	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
89	AF161493	Homo sapiens	HSPC144	1185	100
90	AF161493	Homo sapiens	HSPC144	856	100
91	B25780	787	Human secreted protein SEQ ID	647	41
92	U57344	Mus musculus	Meis3	1007	89
93	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1197	98
94	AL390114	Leishmania major	extremely cysteine/valine rich protein	223	29
95	AB016886	Arabidopsis thaliana	contains similarity to adenylate kinase~gene_id:MCA23.18	287	38
96	AC005525	Homo sapiens	F22162_1	1855	96
97	B20997	Homo sapiens	Human nucleic acid-binding protein, NuABP-1.	3836	99
98	AJ006692	Homo sapiens	ultra high sulfer keratin	507	70
99	AF172264	Homo sapiens	Traf2 and NCK interacting kinase, splice variant 1	6942	99
100	L11239	Homo sapiens	homeobox protein	717	100
101	AC004890	Homo sapiens	similar to zinc finger proteins; similar to AAC01956 (PID:g2843171)	2154	98
102	AC003682	Homo sapiens	R28830_2	1287	48
103	AF201839	Rattus norvegicus	dynamin IIIbb isoform	4270	95
104	Y79510	Homo sapiens	Human carbohydrate-associated protein CRBAP-6.	1394	100
105	Y79510	Homo sapiens	Human carbohydrate-associated protein CRBAP-6.	1209	90
106	AL096748	Homo sapiens	hypothetical protein	1216	100
108	X97260	Homo sapiens	Metallothionein 2	381	100
109	AL034422	Homo sapiens	dJ1141E15.2 (novel protein)	433	100
110	AF191338	Homo sapiens	anaphase-promoting complex subunit 4	683	100
111	AL021712	Arabidopsis thaliana	putative protein	185	26
112	AF250138	Homo sapiens	small stress protein-like protein HSP22	1063	100
113	AL109976	Homo sapiens	dJ794I6.1.1 (novel protein)	4176	99
114	Y36151	787	Human secreted protein	668	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
115	AF110399	Homo sapiens	elongation factor Ts	1666	100
116	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2052	99
117	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
118	X04085	Homo sapiens	catalase	2846	100
119	AF147717	Homo sapiens	ubiquitin C-terminal hydrolase UCH37	1695	100
120	X73882	Homo sapiens	microtubule associated protein	3801	99
121	AC004882	Homo sapiens	similar to CAA16821 (PID:g3255952)	3223	100
122	M93311	Homo sapiens	metallothionein-III	421	100
123	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	557	94
124	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	222	53
125	AF232009	Homo sapiens	peroxisomal trans 2-enoyl CoA reductase	1565	99
126	AB004906	Ipomoea purpurea	transposase	146	20
127	M60165	Homo sapiens	guanine nucleotide-binding regulatory protein 2	1832	99
128	Y10319	Homo sapiens	carnitine carrier	1592	100
129	U75467	Drosophila melanogaster	Atu	937	36
130	Z21507	Homo sapiens	human elongation factor-1-delta	494	87
131	Z21507	Homo sapiens	human elongation factor-1-delta	938	100
132	Y58633	Homo sapiens	Protein regulating gene expression PRGE-26.	6745	100
133	Y58633	Homo sapiens	Protein regulating gene expression PRGE-26.	4818	95
134	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
135	U72970	Sus scrofa	calcium/calmodulin-dependent protein kinase II isoform gamma-B	2723	99
136	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	450	100
137	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
138	AF155648	Homo sapiens	putative zinc finger protein	5855	92
139	AF144638	Homo sapiens	sphingosine-1-phosphate lyase	2977	100
140	AF152318	Homo sapiens	protocadherin gamma A1	4778	100
141	B08517	Homo sapiens	Amino acid sequence of a beta-tubulin antigen.	5841	100
142	X56667	Homo sapiens	calretinin	1410	99
143	X92763	Homo sapiens	tafazzins	1605	100
144	Y95293	Homo sapiens	Human GEF containing NEK-like kinase substrate sGNK.	4092	99
145	AF226046	Homo sapiens	GK003	1198	100
146	M22877	Homo sapiens	cytochrome c	554	98
147	AJ272212	Homo sapiens	protein serine kinase	2196	100
148	AB026491	Homo sapiens	PICK1	2114	98
149	AB018580	Homo sapiens	hluPGFS	1699	100
150	X91868	Homo sapiens	six1	1509	100
151	AF266505	Mus musculus	pseudouridine synthase 3	2135	84
152	U29170	Drosophila melanogaster	ANON-23D	883	43
153	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99
154	AY009128	Homo sapiens	ISCU2	138	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
155	AF141315	Homo sapiens	alpha-1,4-N-acetylglucosaminyltransferase	1842	100
156	AF110645	Homo sapiens	candidate tumor suppressor p33 ING1 homolog	1294	99
157	AF159297	Zea mays	extensin-like protein	238	25
158	AL133325	Homo sapiens	dJ984P4.3 (Homeobox protein NKX2B)	1437	100
159	AF073298	Homo sapiens	small EDRK-rich factor 2	294	100
160	AC004858	Homo sapiens	U1 small ribonucleoprotein 1SNRP homolog; match to PID:g4050087	4032	100
161	AB012109	Homo sapiens	APC10	990	100
162	AL162751	Arabidopsis thaliana	putative protein	194	32
163	AJ005698	Homo sapiens	poly(A)-specific ribonuclease	3351	100
164	AF117646	Homo sapiens	long CBL-3 protein	2547	99
165	AC004002	Homo sapiens	similar to ciliary dynein beta heavy chain; 78% Similarity to P23098 (PID:g118965)	5065	100
166	M10942	Homo sapiens	human metallothionein-Ie	381	100
167	AF126484	Homo sapiens	CARD4	4961	100
168	AF161518	Homo sapiens	HSPC169	1604	100
169	M64983	Homo sapiens	fibrinogen beta chain	2482	100
170	M64983	Homo sapiens	fibrinogen beta chain	2679	100
171	M58514	Gallus gallus	fibrinogen beta chain	1059	78
172	AF078845	Homo sapiens	16.7Kd protein	786	100
173	AC004774	Homo sapiens	Dlx-6	923	100
174	Z98974	Schizosaccharomyces pombe	putative vacuolar protein sorting-associated protein	185	31
175	X56203	Plasmodium falciparum	liver stage antigen	283	23
176	W74726	Homo sapiens	Human secreted protein fg949_3.	1879	100
177	AJ222967	Homo sapiens	cystinosin	1920	100
178	AC024796	Caenorhabditis elegans	contains similarity to TR:O76167	221	27
179	Y66632	Homo sapiens	Membrane-bound protein PRO276.	1370	100
180	AF151803	Homo sapiens	CGI-45 protein	215	28
181	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
182	Y17292	Homo sapiens	Human cell death preventing kinase (DPK-1) protein sequence.	2676	100
183	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	148	27
184	AF151855	Homo sapiens	CGI-97 protein	1214	96
185	AF289664	Mus musculus	CYLN2	4673	90
186	AL022238	Homo sapiens	dJ1042K10.2 (supported by GENSCAN, FGENES and GENEWISE)	4059	100
187	AL022238	Homo sapiens	dJ1042K10.2 (supported by GENSCAN, FGENES and GENEWISE)	2332	100
188	X83543	Homo sapiens	APXL	8513	99
189	AF059569	Homo sapiens	actin binding protein MAYVEN	3106	99
190	M18135	Rattus norvegicus	smooth-muscle alpha tropomyosin	1306	95
191	AF242194	Drosophila melanogaster	brakeless-B	147	52
192	D30689	Bacillus subtilis	subunit of nitrite reductase	113	29
193	Y44984	Homo sapiens	Human epidermal protein-1.	538	97

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
194	B25679	Homo sapiens	Human secreted protein sequence encoded by gene 15 SEQ ID NO:68.	760	100
195	AB020315	787	homologue of mouse dkk-1 gene:Acc	1466	100
196	U35730	Mus musculus	jerky	2021	75
197	AL136450	Homo sapiens	dJ510O21.1 (novel protein)	632	100
198	X56203	Plasmodium falciparum	liver stage antigen	512	24
199	Y70775	Homo sapiens	Follistatin-related protein zfsta.	2027	63
200	X87237	Homo sapiens	a-glucosidase I	4447	99
201	AF101078	Caenorhabditis elegans	CLU-1	1393	46
202	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6611	100
203	X00474	Homo sapiens	pS2 precursor	466	100
204	AB029333	Halocynthia roretzi	HrPET-1	974	54
205	AF146019	Homo sapiens	hepatocellular carcinoma antigen gene 520	998	100
206	AF071002	Homo sapiens	minK-related peptide 1; MiRP1	632	100
207	AB038162	Homo sapiens	trefoil factor 2	744	100
208	U30521	Homo sapiens	P311 HUM	363	100
209	AB000911	Sus scrofa	ribosomal protein	782	100
210	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	3545	100
211	AF180920	Homo sapiens	cyclin L ania-6a	2722	100
212	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5624	100
213	U29244	Caenorhabditis elegans	similar to human (TRE) transforming protein (PIR:S22157)	602	32
214	AL033538	Homo sapiens	dJ477H23.1 (novel protein)	3195	100
215	X52011	Homo sapiens	muscle determination factor	1262	100
216	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
217	AF006751	Homo sapiens	ES/130	4793	99
218	AB007859	Homo sapiens	KIAA0399 protein	3559	99
219	AK026291	Homo sapiens	unnamed protein product	826	100
221	Y84045	Homo sapiens	Splice variant of cancer associated polypeptide CH1-9a11-2.	5851	97
222	Z67996	Homo sapiens	tenascin-R (restrictin)	7186	100
223	AF134802	Homo sapiens	cofilin isoform 1	846	100
224	Y17711	Homo sapiens	atopy related autoantigen CALC	1611	99
225	AF190051	Gallus gallus	hepatocyte nuclear factor 1a dimerization cofactor isoform	443	81
226	AK026256	Homo sapiens	unnamed protein product	866	98
227	Z69368	Schizosacchar omyces pombe	nuf2-like coiled-coil protein	230	25
228	AF275948	Homo sapiens	ABCA1	11763	99
229	AF161384	Homo sapiens	HSPC266	2006	98
230	Y16270	Homo sapiens	paralemin	1951	100
231	AJ245599	Homo sapiens	putative secreted ligand	2379	99
232	W88499	Homo sapiens	Human stomach carcinoma clone HP10412-encoded protein.	1545	99
233	AF096286	Mus musculus	pecanex 1	3623	93
234	V64619_cd 1	Homo sapiens	30-NOV-1990 Human HE1 cDNA.	796	100
235	V64619_cd 1	Homo sapiens	30-NOV-1990 Human HE1 cDNA.	470	98
236	AF227258	Bos taurus	RPGR-interacting protein-1	1262	38
237	AJ132445	Homo sapiens	claudin-14	1181	100
238	AL034562	Homo sapiens	dJ684O24.2 (prodynorphin (Beta-	1330	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			Neoendorphin-Dynorphin precursor, Proenkephalin B precursor))		
239	AF262027	Homo sapiens	eIF-5A2	808	100
240	AL079344	Arabidopsis thaliana	putative protein	194	33
241	AC002394	Homo sapiens	Gene product with similarity to dynein beta subunit	1542	51
242	AJ271361	Takifugu rubripes	FRANK2 protein	303	30
243	AL021918	Homo sapiens	b34I8.1 (Kruppel related Zinc Finger protein 184)	1476	48
244	AF190167	Homo sapiens	membrane associated protein SLP-2	1736	99
245	Y10601	Homo sapiens	ankyrin-like protein	5877	100
246	AL121771	Homo sapiens	dJ548G19.1.1 (novel protein (ortholog of mouse zinc finger protein ZFP64) (translation of cDNA NT2RP3001398 (Em:AK001596)) (isoform 1))	3628	100
247	L25314	Drosophila melanogaster	actin-related protein	984	47
248	X63745	Homo sapiens	KDEL receptor	1095	100
249	AF112208	Homo sapiens	13kDa differentiation-associated protein	816	100
250	AP001707	Homo sapiens	human gene for claudin-8, Accession No. AJ250711	1172	100
251	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	778	100
252	AL031186	Homo sapiens	bK984G1.1 (supported by FGENES)	532	100
253	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	639	100
254	AL049843	Homo sapiens	dJ392M17.3 (KIAA0349 protein)	6741	99
255	AJ242972	Homo sapiens	TOLLIP protein	1424	99
256	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
257	AF279865	Homo sapiens	kinesin-like protein GAKIN	2903	100
258	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
259	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
260	AF101784	Homo sapiens	b-TRCP variant E3RS-IkappaB	3226	99
261	AF101784	Homo sapiens	b-TRCP variant E3RS-IkappaB	2821	100
262	AF101784	Homo sapiens	b-TRCP variant E3RS-IkappaB	3149	99
263	AF197060	Homo sapiens	src homology 3 domain-containing protein HIP-55	2257	100
264	Y86262	Homo sapiens	Human secreted protein HAQAR23, SEQ ID NO:177.	766	100
265	Y56966	Homo sapiens	Human SBPSAPL polypeptide.	2779	100
266	Y56966	Homo sapiens	Human SBPSAPL polypeptide.	1018	99
267	AJ300465	Homo sapiens	putative white family ATP-binding cassette transporter	1557	95
268	AC004030	Homo sapiens	F21856_2	3579	99
269	X55954	Homo sapiens	HL23 ribosomal protein	714	100
270	AB033921	Mus musculus	Ndr1 related protein Ndr2	1855	94
271	AF081886	Homo sapiens	ERO1-like protein	1905	99
272	AF166492	Homo sapiens	small GTPase RAB6B	1060	100
273	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
274	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1530	99
275	X00129	Homo sapiens	precursor RBP	1044	97
276	Z47500_cd1	Homo sapiens	11-MAY-1998 Human RHOH gene sequence.	1161	100
277	AB049188	Equus caballus	ubiquitin C-terminal hydrolase	1118	96

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
278	AF270647	Homo sapiens	GTT1	1564	100
279	AF143956	Mus musculus	coronin-2	2414	94
280	R85151	Homo sapiens	Endothelial cell polypeptide.	911	92
281	R85151	Homo sapiens	Endothelial cell polypeptide.	1031	100
282	D83948	Rattus norvegicus	S1-1 protein	3975	90
283	Y14768	Homo sapiens	I Kappa B-like protein	2037	100
286	AL031316	Homo sapiens	dJ28O10.3(HSD11B1 (hydroxysteroid (11-beta) dehydrogenase 1)	294	100
287	D64109	Homo sapiens	tob family	1773	99
288	AB026043	Homo sapiens	MS4A7	1230	100
289	M61866	Homo sapiens	Krueppel-related DNA-binding protein	209	90
290	AJ001810	Homo sapiens	mRNA cleavage factor I 25 kDa subunit	1217	100
291	Y99454	Homo sapiens	Human PRO1605 (UNQ786) amino acid sequence SEQ ID NO:395.	694	100
292	Y44824	Homo sapiens	Human molecule associated with cell proliferation, MACP-4.	2370	100
293	AJ276101	Homo sapiens	GPRC5B protein	2099	100
294	AF161406	Homo sapiens	HSPC288	719	100
295	Y58628	Homo sapiens	Protein regulating gene expression PRGE-21.	1276	100
296	U91561	Rattus norvegicus	pyridoxine 5'-phosphate oxidase	1239	87
297	L02956	Xenopus laevis	ribonucleoprotein	1624	83
298	AF226730	Homo sapiens	Cyt19	1729	99
299	AF226730	Homo sapiens	Cyt19	906	98
300	Y54324	Homo sapiens	Amino acid sequence of a human gastric cancer antigen protein.	718	89
301	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform	1606	100
302	Y32206	Homo sapiens	Human receptor molecule (REC) encoded by Incyte clone 2825826.	1676	98
303	AF247565	Homo sapiens	hepatocellular carcinoma associated ring finger protein	525	100
304	AF208844	Homo sapiens	BM-002	428	100
305	AC004983	Homo sapiens	similar to PID:g3877944	1988	100
306	AL132978	Arabidopsis thaliana	putative protein	210	25
307	Y10530	Homo sapiens	olfactory receptor	1645	100
308	AF180681	Homo sapiens	guanine nucleotide exchange factor	3597	100
309	AF111856	Homo sapiens	sodium dependent phosphate transporter isoform NaPi-3b	3591	99
310	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
311	Z73420	Homo sapiens	cE146D10.2 (mercaptopyruvate sulfurtransferase (EC 2.8.1.2))	1598	100
312	X79535	Homo sapiens	beta tubulin	2348	100
313	AF070658	Homo sapiens	HSPC002	861	100
314	AF078866	Homo sapiens	SURF-4	1395	100
317	Z37986	Homo sapiens	phenylalkylamine binding protein	1258	100
320	AB047892	Macaca fascicularis	hypothetical protein	258	82
321	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1440	100
322	AB016531	Homo sapiens	PEX16	1741	100
323	AL391141	Arabidopsis	putative protein	274	49

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		thaliana			
325	AF140501	Homo sapiens	DNA polymerase iota	3691	99
326	X96698	Homo sapiens	D1075-like	1450	96
327	AF152325	Homo sapiens	protocadherin gamma A5	4769	100
328	AF151803	Homo sapiens	CGI-45 protein	1970	100
329	X74070	Homo sapiens	transcription factor BTF3	639	81
330	AF171102	Homo sapiens	retinal degeneration B beta	1302	95
331	W54040	Homo sapiens	Human interferon-inducible protein, HIFI.	484	98
332	AF024617	Homo sapiens	transcription-associated zinc ribbon protein	691	100
333	U19181	Rattus norvegicus	Rabin3	2129	90
334	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
335	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
336	AF110774	Homo sapiens	adrenal gland protein AD-001	647	100
337	AB011414	Homo sapiens	Kruppel-type zinc finger protein	1674	58
338	AF207600	Homo sapiens	ethanolamine kinase	129	100
340	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	50
341	Y28576	Homo sapiens	Secreted peptide clone pe503_1.	944	100
342	U32274	Saccharomyces cerevisiae	Ydr386wp; CAI: 0.12	191	37
343	A01771	synthetic construct	vascular anticoagulating protein	1661	99
344	AF220052	Homo sapiens	uncharacterized hematopoietic stem/progenitor cells protein MDS032	1285	100
345	Y70400	Homo sapiens	Human cell-signalling protein-2.	754	100
346	Y50926	Homo sapiens	Human fetal brain cDNA clone vc16_1 derived protein.	962	100
347	AF183428	Homo sapiens	28.4 kDa protein	1329	100
348	AC006069	Arabidopsis thaliana	putative cleavage and polyadenylation specifity factor	1383	55
349	AL032631	Caenorhabditis elegans	Y106G6H.8	194	39
350	U70669	Homo sapiens	Fas-ligand associated factor 3	167	23
351	Y93468	Homo sapiens	Amino acid sequence of a potassium channel interactor protein.	1182	92
352	AF005856	Drosophila yakuba	anon2A5	111	45
353	AJ271684	Homo sapiens	myeloid DAP12-associating lectin	1013	100
354	AF099100	Homo sapiens	WD-repeat protein 6	2882	99
355	U51730	Murine leukemia virus	reverse transcriptase	316	42
356	D50617	Saccharomyces cerevisiae	YFL042C	279	27
357	D50617	Saccharomyces cerevisiae	YFL042C	279	27
358	AF161432	Homo sapiens	HSPC314	1059	93
359	AB029488	Homo sapiens	C11orf21	758	99
360	AJ251024	Homo sapiens	putative odorant binding protein ag	1239	100
361	U43281	Saccharomyces cerevisiae	Lpg22p	2074	74
362	U43281	Saccharomyces cerevisiae	Lpg22p	2153	74

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
363	AC007153	Arabidopsis thaliana	100632	156	24
364	AF197927	Homo sapiens	AF5q31 protein	3992	99
365	D28500	Homo sapiens	mitochondrial isoleucine tRNA synthetase	4286	98
366	X97868	Homo sapiens	arylsulphatase	3141	98
367	AL162048	Homo sapiens	hypothetical protein	1532	100
368	L36062	Mus musculus	steroidogenic acute regulatory protein	189	25
369	AF113249	Homo sapiens	multiple domain putative nuclear protein	1022	59
370	M15888	Bos taurus	endozepine-related protein precursor	2425	84
371	X66363	Homo sapiens	serine/threonine protein kinase	2562	100
372	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1532	89
373	AF100772	Homo sapiens	tenascin-M1	11535	99
374	AF090934	Homo sapiens	PRO0518	382	100
375	AB021643	Homo sapiens	gonadotropin inducible transcription repressor-3	2761	99
376	AB049758	Homo sapiens	MAWD binding protein	1331	100
377	AF070666	Homo sapiens	Kruppel-associated box protein	466	97
378	S59342	Mus sp.	nuclear pore complex glycoprotein p62	464	60
379	AF149205	Mus musculus	Su(var)3-9 homolog Suv39h2	1690	88
380	AF227906	Homo sapiens	UDP-glucose:glycoprotein glucosyltransferase 2 precursor	7851	99
381	AF118566	Mus musculus	hematopoietic zinc finger protein	1769	92
382	AK000619	Homo sapiens	unnamed protein product	810	100
383	AF227906	Homo sapiens	UDP-glucose:glycoprotein glucosyltransferase 2 precursor	7851	99
384	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	2363	100
385	AF125390	Drosophila melanogaster	L82G	139	41
386	Y94907	Homo sapiens	Human secreted protein clone ca106_19x protein sequence SEQ ID NO:20.	1092	50
387	U18795	Saccharomyce s cerevisiae	Yel064cp	206	28
388	AF177388	Homo sapiens	cancer-amplified transcriptional coactivator ASC-2	10748	99
389	AJ002744	Homo sapiens	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase 7	3469	96
390	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger	3166	100
391	AF217525	Homo sapiens	Down syndrome cell adhesion molecule	5337	60
392	U81035	Rattus norvegicus	ankyrin binding cell adhesion molecule neurofascin	3967	91
393	X65224	Gallus gallus	neurofascin	4097	78
394	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	4292	99
395	AF151083	Homo sapiens	HSPC249	444	98
396	AB017026	Mus musculus	oxysterol-binding protein	2173	98
397	AL035587	Homo sapiens	dJ475N16.4 (KIAA0240)	2393	100
398	W74813	Homo sapiens	Human secreted protein encoded by gene 85 clone HSDFV29.	722	92
399	Y71110	Homo sapiens	Human Hydrolase protein-8 (HYDRL-8).	1637	99

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
400	AF039718	Caenorhabditis elegans	contains similarity to lupus LA protein homologs	325	43
401	AE000877	Methanothermobacter thermoautotrophicus	conserved protein	231	36
402	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
403	Z50853	Homo sapiens	CLPP	615	100
405	X03475	Rattus norvegicus	ribosomal protein L35a (aa 1-110)	576	99
406	AF144237	Homo sapiens	LOMP protein	252	44
407	U20239	Mus musculus	fibrosin	288	76
409	AL033378	Homo sapiens	dJ323M4.1 (KIAA0790 protein)	6026	99
410	X54326	Homo sapiens	glutaminyl-tRNA synthetase	7577	99
411	X61585	Bos taurus	polynucleotide adenylyltransferase	3715	97
412	AF217190	Homo sapiens	MLEL1 protein	5271	99
414	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
415	AJ245922	Homo sapiens	alpha-tubulin 8	2370	100
416	AF203032	Homo sapiens	neurofilament protein	220	21
417	Z97653	Homo sapiens	c380A1.2.1 (novel protein (isoform 1))	1567	100
418	AJ404326	Homo sapiens	SR+89	1871	99
419	AJ404326	Homo sapiens	SR+89	902	64
420	AF134726	Homo sapiens	G9A	5334	99
421	L28125	Podospora anserina	beta transducin-like protein	288	39
422	W21733	Homo sapiens	NIP-1 encoded by clone 59.	110	72
423	S67970	Homo sapiens	ZNF75=KRAB zinc finger	951	76
424	L28035	Mus musculus	protein kinase C gamma	3768	98
426	Y73373	Homo sapiens	HTRM clone 921803 protein sequence.	555	56
427	Y73373	Homo sapiens	HTRM clone 921803 protein sequence.	266	49
428	X61118	Homo sapiens	TTG-2a/RBTN-2a	876	100
429	Z96932	Homo sapiens	nuclear autoantigen fo 14 kDa	496	83
430	AJ277291	Homo sapiens	HELG protein	678	72
431	X82157	Homo sapiens	hevin	3525	99
432	AC007192	Homo sapiens	P85B_HUMAN; PTDINS-3-KINASE P85-BETA	3825	99
433	AL021918	Homo sapiens	b34I8.1 (Kruppel related Zinc Finger protein 184)	1713	50
434	AF084464	Rattus norvegicus	GTP-binding protein REM2	141	29
435	AL049795	Homo sapiens	dJ622L5.2 (novel protein)	1756	98
436	M14513	Rattus norvegicus	(Na+ and K+) ATPase, alpha(III) catalytic subunit	4269	99
437	U33460	Homo sapiens	DNA-directed RNA polymerase I, largest subunit	8777	98
438	D87076	Homo sapiens	similar to human bromodomain protein BR140(JC2069)	3067	100
439	L43912	Macaca mulatta	mannose-binding protein A	589	93
440	D31763	Homo sapiens	ha0946 protein is Kruppel-related.	927	49
441	U70976	Homo sapiens	arrestin	2068	99
442	B08069	Homo sapiens	A human beta-alanine-pyruvate aminotransferase (HAPA).	2343	99
443	AF100662	Caenorhabditis	contains similarity to ubiquitin	166	24

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		elegans	carboxyl-terminal hydrolase (Pfam: UCH-1.hmm, score: 28.46) (Pfam: UCH-2.hmm, score: 47.53)		
444	D78017	Rattus norvegicus	NFI-A1	2667	98
445	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2418	100
448	AJ242540	Volvox carteri f. nagariensis	hydroxyproline-rich glycoprotein DZ-HRGP	165	34
449	AJ133352	Homo sapiens	ZNF237 protein	2006	100
450	AJ133352	Homo sapiens	ZNF237 protein	1025	96
451	AF170708	Homo sapiens	T-box protein TBX3	3700	99
452	AK002080	Homo sapiens	unnamed protein product	1546	99
453	L32977	Homo sapiens	Rieske Fe-S protein	1239	93
454	X51760	Homo sapiens	zinc finger protein (583 AA)	1533	57
455	Y01141	Homo sapiens	Secreted protein encoded by gene 7 clone HTLFA90.	1453	99
456	AB006631	Homo sapiens	The human homolog of mouse Cux-2	6559	100
457	AF067165	Homo sapiens	zinc finger protein 3	977	64
458	AF038169	Homo sapiens	unknown	154	38
459	W75214	Homo sapiens	Human secreted protein encoded by gene 19 clone HRSMC69.	1180	95
460	U97002	Caenorhabditis elegans	similar to acyl-CoA dehydrogenases and epoxide hydrolases; Pfam domain PF00441 (Acyl-CoA_dh), Score=57.4, E-value=1.7e-16, N=2; contains similarity to Pfam domain PF00702 (Hydrolase), Score=57.4, E-value=1e-13, N=1	583	37
461	AK023114	Homo sapiens	unnamed protein product	1041	99
462	M93134	Friend murine leukemia virus	pol protein	289	44
463	AF055473	Homo sapiens	GAGE-8	232	47
466	Y51415	Homo sapiens	Human wild type pKe83 protein.	2625	100
467	Y51417	787	Human pKe83 splice variant protein	2433	100
468	Y57936	Homo sapiens	Human transmembrane protein HTMPN-60.	1629	96
469	D38552	Homo sapiens	The ha1539 protein is related to cyclophilin.	2995	100
470	Y70013	Homo sapiens	Human Protease and associated protein-7 (PPRG-7).	3530	100
471	AJ224747	Homo sapiens	C-terminal variant of hINADL including 2 amino acid exchanges and an insertion of 28 amino acids in frame.	7969	100
472	W99665	Homo sapiens	Human secreted protein clone du157_12 protein.	1546	100
473	W99665	Homo sapiens	Human secreted protein clone du157_12 protein.	998	98
474	X63526	Homo sapiens	homologue to elongation factor 1-gamma from A.salina	2273	99
475	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
476	M60832	Homo sapiens	alpha-2 type VIII collagen	3581	99
477	AF039697	Homo sapiens	antigen NY-CO-31	1213	97
478	AF156929	Sus scrofa	inflammatory response protein 6	1588	83
479	AF264717	Homo sapiens	FYVE domain-containing dual specificity protein phosphatase FYVE-DSP2	5610	99
480	AF044578	Homo sapiens	putative DNA polymerase; POL4P	2478	94
481	X89750	Homo sapiens	TGIF protein	1413	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
482	M93107	Homo sapiens	(R)-3-hydroxybutyrate dehydrogenase	1663	96
483	U58334	Homo sapiens	Bbp/53BP2	1556	41
484	AF151538	Homo sapiens	deoxycytidyl transferase; Rev1p	4281	99
485	Z98884	Homo sapiens	dJ467L1.1 (KIAA0833)	699	73
486	AJ243874	Homo sapiens	oligophrenin-4	3682	100
487	Z11737	Homo sapiens	flavin-containing monooxygenase 4	2969	100
488	X56123	Mus musculus	talin	4353	77
489	AJ278112	Homo sapiens	putative cell cycle control protein	335	23
490	W74843	Homo sapiens	Human secreted protein encoded by gene 115 clone HOVBA03.	1013	98
491	Y41337	Homo sapiens	Human secreted protein encoded by gene 30 clone HRDDV47.	509	36
492	X90530	Homo sapiens	ragB	1926	99
493	X90530	Homo sapiens	ragB	1405	99
494	X90530	Homo sapiens	ragB	1893	96
495	AL022394	Homo sapiens	dJ511B24.3 (KIAA0395 (probable homeobox protein))	4990	99
496	Y11395	Homo sapiens	lanthionine synthetase C-like protein 1	2168	100
497	AJ010119	Homo sapiens	Ribosomal protein kinase B (RSK-B)	4001	100
498	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
499	X54131	Homo sapiens	protein-tyrosine phosphatase	10465	99
500	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
501	AC004142	Homo sapiens	similar to murine leucine-rich repeat protein; possible role in neural development by protein-protein interactions; 93% similarity to D49802 (PID:g1369906)	3676	100
502	AL117544	Homo sapiens	hypothetical protein	1226	100
503	AF203032	Homo sapiens	neurofilament protein	5115	99
504	AL034417	Homo sapiens	bK215D11.2 (similar to rat gene 33)	2476	100
505	X69090	Homo sapiens	190kD protein	7546	99
506	U58755	Caenorhabditis elegans	coded for by C. elegans cDNA yk34b1.5; coded for by C. elegans cDNA yk13h10.5; coded for by C. elegans cDNA yk46e8.5; coded for by C. elegans cDNA yk46d5.5; coded for by C. elegans cDNA yk43c2.5; coded for by C. elegans cDNA yk46e8.3; coded for by C. elegans cDNA yk43c2.3; coded for by C. elegans cDNA yk46d5.3; coded for by C. elegans cDNA yk13f10.3; coded for by C. elegans cDNA yk34b1.3	782	55
507	AJ293309	Homo sapiens	NHP2 protein	801	100
508	U39045	Rattus norvegicus	cytoplasmic dynein intermediate chain 2B	3241	97
509	AF063231	Mus musculus	cytoplasmic dynein intermediate chain 2	3159	97
510	AF202893	Mus musculus	Kif21b	4336	95
511	Y13115	Homo sapiens	serine/threonine protein kinase	5071	99
512	AB030207	Homo sapiens	G gamma subunit	364	100
513	AF039571	Homo sapiens	peripheral benzodiazepine receptor interacting protein; PBR-IP/PRAX1	495	33
514	AB037883	Homo sapiens	Gb3/CD77 synthase	1916	99

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
515	D90868	Escherichia coli	similar to	1489	100
516	X98834	Homo sapiens	zinc finger protein Hsal2	5290	100
517	AF055668	Mus musculus	apoptosis-linked gene 4, deltaC form	2904	78
518	AF019926	Mus musculus	protein kinase	1694	90
519	M34513	Homo sapiens	omega protein	317	91
520	Y08612	Homo sapiens	88kDa nuclear pore complex protein	2313	99
521	Y08612	Homo sapiens	88kDa nuclear pore complex protein	1561	99
522	AL096766	Homo sapiens	dA59H18.1 (KIAA0767 protein)	2497	100
523	AF186249	Homo sapiens	six transmembrane epithelial antigen of prostate	1790	100
524	AB029012	Homo sapiens	KIAA1089 protein	4933	100
525	AB026893	Homo sapiens	vascular cadherin-2	5962	100
526	X74331	Homo sapiens	DNA primase (p58 subunit)	1720	100
528	AC007228	Homo sapiens	R31665_2	1488	47
529	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2639	100
530	U80446	Caenorhabditis elegans	coded for by C. elegans cDNA yk172e6.3; coded for by C. elegans cDNA yk158f7.3; coded for by C. elegans cDNA yk158f7.5; coded for by C. elegans cDNA yk172e6.5	420	39
531	S76838	Mus sp.	Dbs	4821	88
532	Z82215	Homo sapiens	dJ68O2.2 (myosin, heavy polypeptide 9, non-muscle)	9828	100
533	AF245505	Homo sapiens	adlican	277	31
534	AF300612	Homo sapiens	N-acetylgalactosamine-4-O-sulfotransferase	993	59
535	AL121928	Homo sapiens	bA18I14.3 (pleckstrin and Sec7 domain protein)	3333	99
536	AJ271055	Mus musculus	iroquois homeobox protein 6	1724	76
537	AF180473	Homo sapiens	Not2p	2267	100
538	AF071059	Mus musculus	zinc finger RNA binding protein	1089	51
539	AF023453	Homo sapiens	actin-related protein 3-beta	2219	100
540	AC003030	Homo sapiens	R29828_1	1401	70
541	AC003030	Homo sapiens	R29828_1	2294	100
542	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in AL023803))	2152	100
543	AB006135	Rattus norvegicus	db83	1238	98
544	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
545	Y07595	Homo sapiens	transcription factor TFIIH	2373	100
546	AL133545	Homo sapiens	bA386N14.1 (novel protein similar to a dual specificity phosphatase)	964	99
547	X83618	Homo sapiens	hydroxymethylglutaryl-CoA synthase	2647	100
548	AF134726	Homo sapiens	NG37	4359	99
549	AB035356	Homo sapiens	neurexin I-alpha protein	6948	99
551	AB037901	Homo sapiens	gene amplified in squamous cell carcinoma-1	5215	99
552	AB043634	Homo sapiens	PAR-6A	885	100
553	AP000693	Homo sapiens	partial CDS	4875	99
554	AF002223	Homo sapiens	myotubularin related 1	3490	100
555	AC004893	Homo sapiens	similar to NEDD-4 (KIA0093); similar to P46934 (PID:g1171682)	1611	100
556	AJ404468	Homo sapiens	axonemal dynein heavy chain	8328	100
557	AJ404468	Homo sapiens	axonemal dynein heavy chain	11137	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
558	X65873	Homo sapiens	kinesin heavy chain	4860	100
559	AJ277365	Homo sapiens	polyglutamine-containing protein	592	36
560	AF205600	Homo sapiens	transposase-like protein	407	27
561	X71125	Homo sapiens	glutaminyl-peptide cyclotransferase	1914	100
562	X71125	Homo sapiens	glutaminyl-peptide cyclotransferase	1456	97
563	X54304	Homo sapiens	myosin regulatory light chain	897	100
564	AF250842	Drosophila melanogaster	multiple asters	130	23
565	Y58608	Homo sapiens	Protein regulating gene expression PRGE-1.	1619	99
566	AL121893	Homo sapiens	bA189K21.5 (novel protein similar to retinoblastoma binding protein (RBBP9))	1012	100
567	AL117352	Homo sapiens	dJ876B10.2 (novel protein (ortholog of rat EXO84))	3713	99
568	AF228603	Homo sapiens	pleckstrin 2	1841	100
569	AF239243	Homo sapiens	histone deacetylase 7	3244	86
570	AF087695	Mus musculus	veli 3	989	100
571	AB046381	Homo sapiens	testis-abundant finger protein	1346	99
572	AC005551	Homo sapiens	R26529_2, partial CDS	1020	100
573	Y90290	Homo sapiens	Human peptidase, HPEP-7 protein sequence.	274	52
574	W76734	Homo sapiens	Human mDia Rho targeting protein.	712	32
575	AL121935	Homo sapiens	bA517H2.3 (t-complex 10 (a murine tcp homolog))	853	78
576	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID NO:132.	2123	99
577	AL121716	Homo sapiens	dJ202D23.2 (novel protein)	6329	99
578	AL121716	Homo sapiens	dJ202D23.2 (novel protein)	6329	99
579	X92715	Homo sapiens	KRAB /C2H2 zinc finger protein	3102	97
580	X54637	Homo sapiens	protein tyrosine kinase	5564	98
581	X78817	Homo sapiens	p115	1148	44
582	AJ251245	Rattus norvegicus	SECIS binding protein 2	3086	71
583	AF113125	Homo sapiens	E-1 enzyme	581	100
584	M19529	Sus scrofa	follistatin A	1906	98
585	AF169677	Homo sapiens	leucine-rich repeat transmembrane protein FLRT3	3403	100
586	D87685	Homo sapiens	similar to human transcription factor TFIIS (S34159).	8083	99
587	Y00876	Homo sapiens	Human LAPH-1 protein sequence.	2110	100
588	Y99674	Homo sapiens	Human GTPase associated protein-25.	2111	99
589	D86973	Homo sapiens	similar to Yeast translation activator GCN1 (P1:A48126)	12033	99
590	AL034452	Homo sapiens	dJ682J15.1 (novel Collagen triple helix repeat containing protein)	1979	100
591	Y57396	Homo sapiens	Human lysoenzyme LYC4 polypeptide.	814	100
592	AJ297743	Mus musculus	torsinB protein	1448	85
593	AF164796	Homo sapiens	NADH:ubiquinone oxidoreductase MLRQ subunit homolog	469	100
594	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	749	94
595	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone HLDRM43.	824	100
596	Y77123	Homo sapiens	Human neurotransmission-associated protein (NTAP) 998868.	2102	98
597	AF215703	Drosophila	KISMET-L long isoform	1880	65

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
		melanogaster			
598	AF070447	Homo sapiens	barrier-to-autointegration factor	290	90
599	X56203	Plasmodium falciparum	liver stage antigen	372	22
600	X79828	Mus musculus	NK10	202	53
601	AB004109	Cricetulus griseus	phosphatidylserine synthase II	2262	92
602	U94988	Mus musculus	Nulp1	2912	89
603	U94988	Mus musculus	Nulp1	2800	86
604	AF006264	Homo sapiens	recombination and sister chromatid cohesion protein homolog	2850	100
605	AF006264	Homo sapiens	recombination and sister chromatid cohesion protein homolog	2530	100
606	X82260	Homo sapiens	RanGAP1	2929	100
607	X82260	Homo sapiens	RanGAP1	1843	97
608	AF160909	Drosophila melanogaster	BcDNA.LD03471	943	58
610	X74801	Homo sapiens	gamma subunit of CCT chaperonin	2745	99
611	AL031427	Homo sapiens	dJ167A19.1 (novel protein)	1608	100
612	Y71072	Homo sapiens	Human membrane transport protein, MTRP-17.	445	100
613	X16396	Homo sapiens	precursor polypeptide (AA -29 to 315)	1749	100
614	AK000281	Homo sapiens	unnamed protein product	1814	99
615	AB011128	Homo sapiens	KIAA0556 protein	5761	99
616	U19361	Petromyzon marinus	NF-180	205	21
617	AF045555	Homo sapiens	wbscr1	1208	100
618	AF045555	Homo sapiens	wbscr1 alternative spliced product	1318	100
619	U22229	Felis catus	ribosomal protein L41	128	100
620	Y17169	Homo sapiens	A6 related protein	1819	100
621	Y12065	Homo sapiens	hNop56	2956	99
622	AF177758	Homo sapiens	ubiquitin specific protease 16	2998	100
623	AF317425	Homo sapiens	GAC-1	3866	100
624	AL050297	Homo sapiens	hypothetical protein	1227	99
625	AC007204	Homo sapiens	BC273239_1	3398	99
626	Z68747	Homo sapiens	imogen 38	2024	99
627	Z68747	Homo sapiens	imogen 38	1958	97
628	Y70229	Homo sapiens	Human RNA-associated protein-10 (RNAAP-10).	3424	99
629	AF191492	Homo sapiens	nasopharyngeal carcinoma associated gene protein-8	613	100
630	AF119664	Homo sapiens	transcriptional regulator protein HCNGP	1574	100
631	AF119664	Homo sapiens	transcriptional regulator protein HCNGP	1150	89
632	Y17849	Homo sapiens	ganglioside-induced differentiation associated protein 1	1839	98
633	X55740	Homo sapiens	5'-nucleotidase	3012	100
634	AF039688	Homo sapiens	antigen NY-CO-3	931	100
635	AF119662	Homo sapiens	E46 protein	2424	100
636	AB007836	Homo sapiens	Hic-5	2544	100
637	AF077818	Mus musculus	syntrophin-associated serine-threonine protein kinase	2027	44
638	AL035455	Homo sapiens	dJ1018E9.1 (VAMP (vesicle-associated membrane protein)-associated protein B and C)	150	26
639	AF078844	Homo sapiens	hqp0376 protein	416	81

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
640	U28377	Escherichia coli	ORF_f239; was ORF_f191 and ORF_f194 before splice	1198	100
641	AK024442	Homo sapiens	FLJ00032 protein	1677	56
642	U58682	Homo sapiens	ribosomal protein S28	340	100
643	X57432	Rattus rattus	ribosomal protein S2	1520	98
644	AB002348	Homo sapiens	KIAA0350 protein	5186	99
646	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	1178	98
647	AB029482	Mus musculus	JNK-binding protein JNKBP1	4609	81
648	AB009053	Arabidopsis thaliana	contains similarity to isoamyl acetate-hydrolyzing esterase~gene_id:MQB2.25	407	44
650	AC002550	Homo sapiens	Unknown gene product	858	99
651	U26592	Homo sapiens	diabetes mellitus type I autoantigen	253	66
652	X60155	Homo sapiens	zinc finger 41	4349	100
653	X53330	Platynereis dumerilii	H4 protein (AA 1 - 103)	523	100
654	AC003682	Homo sapiens	R27945_2	2558	100
655	X80473	Mus musculus	rab19	596	56
656	J02649	Rattus norvegicus	unknown protein	201	95
657	AC006014	Homo sapiens	similar to RFP transforming protein; similar to P14373 (PID:g132517)	1331	99
658	X92972	Homo sapiens	protein phosphatase 6	1666	100
659	L35269	Homo sapiens	zinc finger protein	2803	99
660	AC003682	Homo sapiens	F18547_1	3184	96
661	X79204	Homo sapiens	ataxin-1	4195	99
662	X17620	Homo sapiens	Nm23 protein	965	99
663	AB015617	Homo sapiens	ELKS	1501	80
664	Z56281	Homo sapiens	interferon regulatory factor 3	2331	100
665	AJ248283	Pyrococcus abyssi	LACTOYLGLUTATHIONE LYASE (EC 4.4.1.5) METHYLGLYOXALASE) (ALDOKETOMUTASE) (GLYOXALASE I).	254	40
666	Z70200	Homo sapiens	U5 snRNP-specific 200kD protein	8819	99
667	Z70200	Homo sapiens	U5 snRNP-specific 200kD protein	8589	97
668	AF153450	Manduca sexta	juvenile hormone esterase binding protein	225	32
669	AF227198	Homo sapiens	CrkRS	7231	99
670	X99586	Homo sapiens	SMT3C protein	441	87
671	Z61589_cd1	Homo sapiens	17-AUG-1998 DNA encoding a human OC-2 protein.	2593	100
672	AJ132702	Mus musculus	ATFa-associated factor	3240	88
673	AF204159	Homo sapiens	potassium large conductance calcium-activated channel beta 3a subunit	1486	100
674	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
675	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
676	AB016839	Homo sapiens	mob1	419	42
677	D86970	Homo sapiens	similar to myosin heavy chain: Containing ATP/GTP-binding site motif A(P-loop)	161	28
678	U83115	Homo sapiens	non-lens beta gamma-crystallin like protein	8569	99
679	AF203687	Homo sapiens	prolactin regulatory element-binding protein	2181	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
680	M27685	Mus musculus	ultra-high sulphur keratin	650	58
681	U04968	Cricetulus griseus	nucleotide excision repair protein	3712	97
682	AF119663	Homo sapiens	G-protein gamma-12 subunit	356	100
683	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	342	100
684	X67699	Homo sapiens	CDw52 antigen	297	100
685	AF022789	Homo sapiens	ubiquitin hydrolyzing enzyme I	1892	100
686	AJ001006	Mus musculus	EMeg32 protein	938	96
687	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
688	AF019661	Mus musculus	zeta proteasome chain; PSMA5	1214	100
689	AF156557	Homo sapiens	stomatin related protein	2036	100
690	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	593	100
691	AF161512	Homo sapiens	HSPC163	738	100
692	AL031115	Homo sapiens	ZXDA, ZXDB (zinc finger X-linked protein)	4298	100
693	L40410	Homo sapiens	thyroid receptor interactor	806	100
694	AC004542	Homo sapiens	OXYSTEROL-BINDING PROTEIN-like; similar to P22059 (PID:g129308)	2533	99
695	AF169411	Rattus norvegicus	PAPIN	4144	52
696	Y58168	Homo sapiens	Human hydrolase homologue HHH-4.	2144	100
697	AF271994	Homo sapiens	dopamine responsive protein DRG-1	1613	100
698	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
699	AL133506	Unknown	/prediction=(method:""genscan"", version:""1.0"", score:""109.13""); /prediction=(method:	825	48
700	Y96870	Homo sapiens	Human goose-type lysozyme (GOLY).	1032	100
701	AC003034	Homo sapiens	Gene with similarity to rat kidney-specific (KS) gene	1190	100
702	AC003034	Homo sapiens	Gene with similarity to rat kidney-specific (KS) gene	937	95
703	AJ242832	Homo sapiens	calpain	3756	100
704	S52624	Homo sapiens	unknown	185	100
705.	AF005081	Homo sapiens	skin-specific protein	652	100
706	Y16793	Homo sapiens	keratin, type I	2232	100
707	Y44985	Homo sapiens	Human epidermal protein-2.	455	69
708	AF113220	Homo sapiens	MSTP040	686	100
709	Y44985	Homo sapiens	Human epidermal protein-2.	408	65
710	Y16132	Homo sapiens	CDT6	1874	100
711	Y68775	Homo sapiens	Amino acid sequence of a human phosphorylation effector PHSP-7.	2407	100
712	X63422	Homo sapiens	H(+)-transporting ATP synthase	209	100
713	AF169968	Mus musculus	DNA binding protein DESRT	1467	79
714	X52563	Bos taurus	permability increasing protein	383	29
715	AJ277739	Homo sapiens	RPB11b1alpha protein	480	98
716	AL135791	Homo sapiens	bA162G10.3 (zinc finger protein)	401	98
717	AF223466	Homo sapiens	HT015 protein	1311	97
719	AF117383	Homo sapiens	placental protein 13; PP13	746	100
720	Z98743	Homo sapiens	dJ181C9.2 (Rho GTPase activating protein 8 (RhoGAP, p50RhoGAP))	324	100
721	AL163815	Arabidopsis thaliana	putative protein	653	61
722	G01436	Homo sapiens	Human secreted protein, SEQ ID	418	96

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			NO: 5517.		
723	AF282919	Mus musculus	Zfp228	349	49
724	AB023191	Homo sapiens	KIAA0974 protein	2953	100
725	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
726	AL021939	Homo sapiens	dJ352A20.2 (aldehyde dehydrogenase family protein)	1764	100
727	AF182426	Rattus norvegicus	arylacetamide deacetylase	791	42
728	Y08565	Homo sapiens	UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase	3331	99
729	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	1652	99
730	AL078606	Arabidopsis thaliana	putative protein	277	55
731	Y73352	Homo sapiens	HTRM clone 1732368 protein sequence.	1720	100
732	AF178432	Homo sapiens	SH3 protein	3302	100
733	Y17832	Human endogenous retrovirus K	env protein	223	34
734	Y28859	Homo sapiens	Human mesoderm induction early response protein ER1.	2067	98
735	U09355	Oryctolagus cuniculus	protein phosphatase 2A1 B gamma subunit	2352	99
736	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	724	99
737	AB027003	Mus musculus	protein phosphatase	378	84
738	AF112200	Homo sapiens	NADH-oxidoreductase B18 subunit	739	100
739	AF112200	Homo sapiens	NADH-oxidoreductase B18 subunit	613	88
740	AF302154	Homo sapiens	SPG protein	6556	100
741	B25681	Homo sapiens	Human secreted protein sequence encoded by gene 17 SEQ ID NO:70.	1410	99
742	L27479	Homo sapiens	X123	1237	99
743	L27479	Homo sapiens	X123	1206	97
744	Y66745	Homo sapiens	Membrane-bound protein PRO1186.	588	99
745	AJ001019	Homo sapiens	ring finger protein	1292	99
746	X68453	Sus scrofa	tubulin-tyrosine ligase	1882	94
747	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
748	AF151069	Homo sapiens	HSPC235	1694	96
749	AF182404	Homo sapiens	mitochondrial uncoupling protein 1	1674	100
750	AL121993	Homo sapiens	dJ776P7.1 (Novel protein)	2500	99
751	AF149825	Homo sapiens	PACSIN3	2253	100
752	AL008635	Homo sapiens	dJ510H16.2 (high-mobility group protein 2-like 1)	3026	99
753	Y57914	Homo sapiens	Human transmembrane protein HTMPN-38.	1124	100
754	AF285109	Homo sapiens	septin 3 isoform B	1766	100
755	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	2371	95
756	Z19585	Homo sapiens	thrombospondin-4	4239	100
757	AP001745	Homo sapiens	similar to zinc finger 5 protein	1857	100
758	AF190664	Mus musculus	LMBR2	555	72
759	AF090326	Mus musculus	AE-1 binding protein AEBP2	1540	97
760	AL096677	Homo sapiens	dJ322G13.3 (novel protein similar to	999	94

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
			bovine and mouse beta-soluble NSF attachment protein (SNAP-beta))		
761	AC003007	Homo sapiens	Unknown gene product (partial)	649	96
762	U66372	Bos taurus	ribosomal protein S29	230	73
764	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	1152	100
765	U88169	Caenorhabditis elegans	similar to molybdoterin biosynthesis MOEB proteins	1204	65
766	AL118506	Homo sapiens	dJ591C20.3.1 (novel DnaJ domain protein, similar to mouse and bovine cysteine string protein)	1091	100
767	AK024693	Homo sapiens	unnamed protein product	3767	100
768	Z11518	Homo sapiens	histidyl-tRNA synthetase	2582	100
769	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	25529	100
770	AC009360	Arabidopsis thaliana	Contains 3 PF|00400 WD40, G-beta repeat domains.	333	33
771	AB037685	Mus musculus	LANP-like protein	1246	91
772	AL161578	Arabidopsis thaliana	putative protein	335	46
773	AL161578	Arabidopsis thaliana	putative protein	333	47
774	AY008271	Homo sapiens	helicase SMARCAD1	5264	99
775	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	1127	96
776	W88853	Homo sapiens	Polypeptide fragment encoded by gene 89.	752	100
777	W88853	Homo sapiens	Polypeptide fragment encoded by gene 89.	752	100
778	W88853	Homo sapiens	Polypeptide fragment encoded by gene 89.	752	100
779	AF196481	Homo sapiens	RING finger protein; FXY2	3644	100
780	AL035427	Homo sapiens	dJ769N13.1 (KIAA0443 protein.)	1609	54
781	AB026187	Homo sapiens	protocadherin-Xa	5244	100
782	B24458	Homo sapiens	Human secreted protein sequence encoded by gene 22 SEQ ID NO:83.	1002	100
783	AB027289	Homo sapiens	cyclin-E binding protein 1	5421	100
784	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	627	100
785	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
786	AJ245820	Homo sapiens	type I transmembrane receptor	4624	100
787	Z48042	Homo sapiens	GPI-anchored protein p137	3340	99
788	AL031782	Homo sapiens	dJ708F5.1 (PUTATIVE novel Collagen alpha 1 LIKE protein)	2739	100
789	AJ131245	Homo sapiens	Sec24B protein	6602	100
790	AF107203	Homo sapiens	ataxin 2-binding protein	2008	100
791	Y14690	Homo sapiens	procollagen alpha 2(V)	600	34
792	AL031055	Homo sapiens	dJ28H20.2 (novel protein)	1267	100
793	Y36194	787	Human secreted protein	2051	99
794	AB028127	Homo sapiens	mannosyltransferase	2138	96
795	AC007228	Homo sapiens	R31665_2	2738	79
796	AL049482	Arabidopsis thaliana	putative protein	436	47
797	AC004528	Homo sapiens	R32184_3	891	91
798	AB037830	Homo sapiens	KIAA1409 protein	7532	100
799	X53793	Homo sapiens	5' half of the product is homologues to Bacillus subtiis SAICAR synthetase, 3' half corresponds to the catalytic subunit of AIR carboxylase	2232	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
800	Y99350	Homo sapiens	Human PRO1378 (UNQ715) amino acid sequence SEQ ID NO:33.	1343	100
801	AB042636	Homo sapiens	junctophilin type3	1225	47
802	AB029324	Rattus norvegicus	TIP120-family protein TIP120B	3916	90
803	AB029324	Rattus norvegicus	TIP120-family protein TIP120B	4961	90
804	AF251040	Homo sapiens	putative nuclear protein	2119	100
805	AB033281	Homo sapiens	F-box and WD-repeats protein beta-TRCP2 isoform C	2879	100
806	U87305	Rattus norvegicus	transmembrane receptor UNC5H1	3257	90
807	AF118889	Rattus norvegicus	b-tomosyn isoform	3155	97
808	AF226993	Rattus norvegicus	selective LIM binding factor	8793	95
809	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	3939	99
810	AL031782	Homo sapiens	dJ708F5.1 (PUTATIVE novel Collagen alpha 1 LIKE protein)	1546	100
811	AC002542	Homo sapiens	similar to C. elegans F11A10.5; 80% similarity to Z68297 (PID:g1130619)	2294	100
812	U83246	Homo sapiens	copine I	606	52
813	AF242552	Gallus gallus	retinovin	945	34
814	X52332	Homo sapiens	zinc finger protein 10	1651	93
815	X52332	Homo sapiens	zinc finger protein 10	2423	99
816	Y09631	Homo sapiens	PIBF1 protein	2935	99
817	X71997	Rattus norvegicus	myosin I	3883	98
818	AY004877	Mus musculus	cytoplasmic dynein heavy chain	11105	98
819	Y27196	Homo sapiens	Human cyclic nucleotide phosphodiester PDE8B(E) amino acid sequence.	3790	100
820	AF081947	Mus musculus	tektin	1134	81
821	AL035106	Homo sapiens	dJ998C11.1 (continues in Em:AL445192 as bA269H4.1)	871	100
822	AF022795	Homo sapiens	TGF beta receptor associated protein-1	385	24
823	AF015770	Mus musculus	radical fringe	1422	82
824	U82695	Homo sapiens	expressed-Xq28STS protein	1444	99
825	X77371	Mesocricetus auratus	COR1	641	78
826	AB014576	Homo sapiens	KIAA0676 protein	296	79
827	AL049733	Homo sapiens	dJ875H3.1 (APK1 antigen)	1584	72
828	AF222980	Homo sapiens	disrupted in Schizophrenia 1 protein	4418	100
829	Z31560	Homo sapiens	sox-2	1683	100
830	AF295773	Homo sapiens	ral guanine nucleotide dissociation stimulator	4717	99
831	AB041926	Homo sapiens	GCK family kinase MINK-2	6866	100
832	L04948	Saccharomyces cerevisiae	mitochondrial transporter protein	338	35
833	AJ007012	Mus musculus	Fish protein	704	94
834	Z34289	Homo sapiens	nucleolar phosphoprotein p130	3455	99
835	U10991	Homo sapiens	G2	8436	98
836	AF230877	Homo sapiens	MIP-T3	2945	99
837	X58288	Homo sapiens	protein-tyrosine phosphatase	7734	99
838	X56958	Homo sapiens	ankyrin (brank-2)	9631	100
839	AC024791	Caenorhabditis elegans	contains similarity to beta-lactamases	370	24

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
840	D83197	Homo sapiens	ankyrin repeat protein	802	99
841	AF053711	Serinus canaria	neurofilament medium subunit	192	31
842	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	990	96
843	U76343	Homo sapiens	GABA transport protein	2992	98
844	Y13645	Homo sapiens	uroplakin II	897	100
845	D21064	Homo sapiens	similar to rat general mitochondrial matrix processing protease mRNA (RATMPP).	2710	99
846	AF192522	Homo sapiens	Niemann-Pick C3 protein; NPC3	7047	100
847	AF192522	Homo sapiens	Niemann-Pick C3 protein; NPC3	5472	100
848	X60489	Homo sapiens	elongation factor-1-beta	1162	100
849	AC007204	Homo sapiens	BC273239_1	2277	67
850	AC003682	Homo sapiens	R28830_1	2401	100
851	AL121583	Homo sapiens	bA358N2.1 (novel protein)	353	61
852	Z48475	Homo sapiens	glucokinase regulator	3155	99
853	Z83844	Homo sapiens	dJ37E16.2 (SH3-domain binding protein 1)	1884	98
854	AF233323	Homo sapiens	Fas-associated phosphatase-1	390	36
855	AF062741	Rattus norvegicus	pyruvate dehydrogenase phosphatase isoenzyme 2	447	80
856	Y11411	Homo sapiens	pristanoyl-CoA oxidase	3595	98
857	M97188	Strongylocentrotus purpuratus	tektin A1	290	46
858	AB001105	Homo sapiens	hippocalcin-like protein 4	995	100
859	AF164791	Homo sapiens	putative 38.3kDa protein	1795	100
860	AF298117	Homo sapiens	homeobox protein OTX2	1477	93
861	AF015264	Rattus norvegicus	golgi peripheral membrane protein p65	1820	81
862	X16901	Homo sapiens	30kb subunit of RAB30 /74	1284	100
863	M12140	Homo sapiens	envelope protein	202	81
864	AF161459	Homo sapiens	HSPC109	815	98
865	AL109983	Homo sapiens	dJ718P11.1.1 (novel class II aminotransferase similar to serine palmotyltransferase (isoform 1))	444	100
866	M77183	Rattus norvegicus	alpha-1-macroglobulin	227	45
867	AF272663	Homo sapiens	gephyrin	3785	100
868	X75285	Mus musculus	fibulin-2	3258	87
869	X82494	Homo sapiens	fibulin-2	3407	99
870	AJ297743	Mus musculus	torsinB protein	169	43
871	AJ278313	Homo sapiens	phospholipase C-beta-1a	6258	99
872	AF073344	Homo sapiens	ubiquitin-specific protease 3	256	43
873	Y91955	Homo sapiens	Human cytoskeleton associated protein 10 (CYSKP-10).	535	100
874	AJ000414	Homo sapiens	Cdc42-interacting protein 4	1136	53
875	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	627	100
876	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2537	98
877	AF182198	Homo sapiens	intersectin 2 long isoform	8764	99
878	L17308	Gossypium hirsutum	proline-rich cell wall protein	192	35
879	AF177169	Homo sapiens	tropomodulin 2	1769	100
880	W03627	Homo sapiens	Human follicle stimulating hormone GPR N-terminal sequence.	210	23

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
881	AL021068	Homo sapiens	dJ206D15.3	2615	99
882	AC005498	Homo sapiens	R31665_2	318	82
883	AF165518	Homo sapiens	MAGOH isoform	182	94
884	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	43
885	U13045	Homo sapiens	nuclear respiratory factor-2 subunit beta 1	869	62
886	X52836	Homo sapiens	tryptophan hydroxylase (AA 1 - 444)	2320	98
887	X51466	Homo sapiens	elongation factor 2	4460	100
888	AB039903	Homo sapiens	interferon-responsive finger protein 1 long form	1096	98
889	X51760	Homo sapiens	zinc finger protein (583 AA)	3130	100
890	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	1024	100
891	W67928	Homo sapiens	Fragment of human secreted protein encoded by gene 4.	391	100
892	AB020598	Homo sapiens	peptide transporter 3	3017	100
893	Y66648	Homo sapiens	Membrane-bound protein PRO1120.	4722	99
894	Y66648	Homo sapiens	Membrane-bound protein PRO1120.	3606	96
895	A29218_cd 1	Homo sapiens	19-NOV-1998 DNA encoding G-protein coupled 7 TM receptor with AXOR15 activity.	2178	100
896	AJ000332	Homo sapiens	Glucosidase II	5063	99
897	X98259	Homo sapiens	M-phase phosphoprotein 8	1085	100
898	X57110	Homo sapiens	c-cbl protein	4849	99
899	X63652	Homo sapiens	inter-alpha-trypsin inhibitor heavy chain ITIH1	3376	98
900	X85134	Homo sapiens	RB protein binding protein	2816	99
901	L11672	Homo sapiens	zinc finger protein	2047	58
902	Y85565	Homo sapiens	Human homologue of UNC-53 (Hs-UNC-53/2) sequence.	369	83
903	X54871	Homo sapiens	ras related protein Rab5b	1094	100
904	Z98265	Homo sapiens	plakophilin 3	4065	100
905	AL035295	Homo sapiens	hypothetical protein	959	99
906	AF051782	Homo sapiens	diaphanous 1	801	35
907	AF208536	Homo sapiens	nucleotide binding protein; NBP	1372	100
908	U79240	Homo sapiens	serine/threonine protein kinase	2365	98
909	U79240	Homo sapiens	serine/threonine protein kinase	2386	99
910	AJ132545	Homo sapiens	protein kinase	2921	100
911	AJ132545	Homo sapiens	protein kinase	1637	99
912	AL121733	Homo sapiens	hypothetical protein	1344	99
913	Y67579	Homo sapiens	Human death inducer-obliterator 1 (DIO-1) polypeptide.	1586	100
914	X87342	Homo sapiens	Human giant larvae homologue	5317	99
915	X87342	Homo sapiens	Human giant larvae homologue	3495	96
916	M94362	Homo sapiens	lamin B2	2357	93
917	AJ011654	Homo sapiens	triple LIM domain protein	3432	100
918	AJ131899	Rattus norvegicus	proline rich synapse associated protein 1	5776	88
919	AF054986	Homo sapiens	putative transmembrane GTPase	1816	100
920	U95822	Homo sapiens	putative transmembrane GTPase	1237	100
921	Y11588	Homo sapiens	apoptosis specific protein	1492	100
922	X84195	Homo sapiens	acylphosphatase	510	100
923	U72882	Homo sapiens	interferon-induced leucine zipper protein	1409	99
924	AE000660	Homo sapiens	hADV36S1	573	100
925	AF126245	Homo sapiens	acyl-Coenzyme A dehydrogenase-8 precursor	2162	100

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
926	AE001968	Deinococcus radiodurans	hypothetical protein	147	27
927	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-2) polypeptide.	1778	100
928	U01317	Homo sapiens	beta-globin	687	94
929	X98333	Homo sapiens	organic cation transporter	2933	100
930	Y91444	Homo sapiens	Human secreted protein sequence encoded by gene 42 SEQ ID NO:165.	1401	100
931	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	1243	100
932	D90279	Homo sapiens	collagen alpha 1(V) chain precursor	569	39
933	Z31560	Homo sapiens	sox-2	1587	96
934	AF147790	Homo sapiens	transmembrane mucin 12	3047	99
935	Z85996	Homo sapiens	match: multiple proteins; match: Q08151 P28185 Q01111 Q43554; match: Q08150 Q40195 P20340 Q39222; match: Q40368 P36412 P40393 Q40723; match: CE01798 Q38923 Q40191 Q41022; match: Q39433 Q40177 Q40218 Q08146; match: P10949 P11023 Q16948 Q20337; match: Q25389 P25228 P20336 P05713; match: P35276 Q08147 P17609 P22128; match: Q15771 P36410 P35291; GTP-binding	726	94
936	AB041533	Homo sapiens	sperm antigen	1054	38
937	X91906	Homo sapiens	voltage-gated chloride ion channel	3914	100
938	AB032481	Homo sapiens	homeobox transcription factor	1744	100
939	AF111106	Homo sapiens	protein serine/threonine phosphatase 4 regulatory subunit 1	4682	99
940	Y17999	Homo sapiens	Dyrk1B protein kinase	3331	99
941	AF305872	Homo sapiens	thyroglobulin	455	92
942	AF263462	Homo sapiens	cingulin	5939	99
943	AK024442	Homo sapiens	FLJ00032 protein	1616	61
944	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 160.	262	35
945	AB015320	Homo sapiens	sigma1B subunit of AP-1 clathrin adaptor complex	599	71
946	Z82287	Caenorhabditis elegans	ZK550.2	229	35
947	D84223	Homo sapiens	leucyl tRNA synthetase	6207	99
948	U49057	Rattus norvegicus	rA9	3846	62
949	AK000568	Homo sapiens	unnamed protein product	1659	100
950	AL021578	Homo sapiens	dJ453C12.6.1 (uncharacterized hypothalamus protein (isoform 1))	257	42
951	AB032435	Homo sapiens	differentiation-associated Na-dependent inorganic phosphate cotransporter	3063	99
952	AF110532	Homo sapiens	uncoupling protein UCP-4	1561	100
953	X83587	Mus musculus	1A13 protein	1420	59
954	AL031665	Homo sapiens	dJ545L17.5.1 (novel protein)	386	53
955	Y87600	Homo sapiens	Human fatty acid synthase-like protein (HFASLP).	2377	100
956	Y99421	Homo sapiens	Human PRO1433 (UNQ738) amino acid sequence SEQ ID NO:292.	522	55

SEQ ID NO:	ACCESSION NUMBER	SPECIES	DESCRIPTION	SMITH-WATERMAN SCORE	% IDENTITY
957	U68535	Mus musculus	aldo-keto reductase	451	73
958	AC007067	Arabidopsis thaliana	T10O24.10	1594	57
959	U72194	Mus musculus	muskelin	3947	99
960	AE003661	Drosophila melanogaster	CG15168 gene product	277	54
961	X80332	Mus musculus	rab20	983	82
962	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
963	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
964	L32602	Rattus norvegicus	homeodomain 159..341	1821	96
965	Z97832	Homo sapiens	dJ329A5.3 (KIAA06460 protein)	3581	99
966	W88995	Homo sapiens	Polypeptide fragment encoded by gene 146.	176	39
967	U12465	Homo sapiens	ribosomal protein L35	604	100
968	AF151803	Homo sapiens	CGI-45 protein	1101	78
969	W74865	Homo sapiens	Human secreted protein encoded by gene 137 clone HMWIF35.	1348	98
970	L21936	Homo sapiens	succinate dehydrogenase flavoprotein subunit	703	100
971	AJ133521	Drosophila buzzatii	protease, reverse transcriptase, ribonuclease H, integrase	194	23
972	AC006017	Homo sapiens	N-acetylgalactosaminyltransferase; similar to Q10473 (PID:g1709559)	3271	100
973	Z81317	Schizosacchar omyces pombe	DNA2-NAM7 helicase family protein	685	31
974	M17885	Homo sapiens	acidic ribosomal phosphoprotein (P0)	792	100
975	U22829	Mus musculus	P2Y purinoceptor	399	40
976	AL132772	Homo sapiens	dJ1013A22.1 (hepatic nuclear factor 4, alpha)	2466	99
977	AC003973	Homo sapiens	ZNF91L	1550	43
978	J04031	Homo sapiens	MDMCSF (EC 1.5.1.5; EC 3.5.4.9; EC 6.3.4.3)	2824	63
979	AF136715	Homo sapiens	taxol resistant associated protein	217	76
980	AF136715	Homo sapiens	taxol resistant associated protein	306	95
981	Z92822	Caenorhabditis elegans	ZK520.1	1109	44
982	AJ295149	Homo sapiens	putative dipeptidase	1564	99
983	AL021331	Homo sapiens	dJ366N23.3 (KIAA0173 and Tubulin-Tyrosine Ligase LIKE)	1492	100
984	AL161501	Arabidopsis thaliana	putative adenosine deaminase	370	38

TABLE 3

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
2	BL00282	Kazal serine protease inhibitors family proteins.	BL00282 16.88 4.259e-14 97-120
3	BL00298	Heat shock hsp90 proteins family proteins.	BL00298A 10.97 1.000e-40 74-119 BL00298E 27.30 1.000e-40 321-376 BL00298F 11.21 1.000e-40 409-464 BL00298H 20.50 1.000e-40 553-607 BL00298C 16.40 2.286e-40 186-230

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00298B 15.64 1.290e-39 134-181 BL00298G 24.57 5.345e-39 465-520 BL00298I 30.07 7.818e-34 661-715 BL00298D 17.97 6.226e-33 242-282
4	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237A 11.48 4.316e-13 57-82
5	PD02454	!!!! PROTEIN ALU SUBFAMILY WARNING ENTRY NUCLEAR PHOSPHO.	PD02454B 11.61 4.309e-17 75-103
6	DM00864	EGF-LIKE DOMAIN.	DM00864A 15.21 7.429e-09 98-119
7	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237A 11.48 1.750e-11 29-54 PR00237D 8.94 7.000e-09 138-160 PR00237B 13.50 8.250e-09 61-83
9	PF00855	PWWP domain proteins.	PF00855 13.75 5.667e-15 272-289
10	BL00139	Eukaryotic thiol (cysteine) proteases cysteine proteins.	BL00139D 9.24 4.400e-11 391-408 BL00139A 10.29 7.511e-09 67-77
12	BL01113	C1q domain proteins.	BL01113B 18.26 9.294e-19 689-725 BL01113C 13.18 4.857e-11 757-777 BL01113D 7.47 2.161e-10 790-800
13	BL01113	C1q domain proteins.	BL01113B 18.26 3.813e-14 599-635 BL01113C 13.18 4.857e-11 667-687 BL01113D 7.47 2.161e-10 700-710
14	BL00594	Aromatic amino acids permeases proteins.	BL00594A 16.75 6.531e-10 50-94
15	BL01047	Heavy-metal-associated domain proteins.	BL01047B 19.73 4.913e-13 707-728
16	PR00625	DNAJ PROTEIN FAMILY SIGNATURE	PR00625A 12.84 7.462e-18 310-330 PR00625B 13.48 3.939e-15 340-361
18	BL00615	C-type lectin domain proteins.	BL00615A 16.68 3.700e-09 144-162
20	PR00741	GLYCOSYL HYDROLASE FAMILY 29 SIGNATURE	PR00741D 16.11 9.082e-21 175-195 PR00741F 14.66 9.262e-21 243-265 PR00741B 14.23 1.947e-18 128-145 PR00741G 9.29 2.180e-17 318-340 PR00741C 9.16 7.328e-17 147-166 PR00741H 10.32 2.141e-13 351-374 PR00741A 9.24 3.596e-13 89-105 PR00741E 13.39 3.535e-12 215-232
22	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.647e-20 117-148 BL00107B 13.31 1.000e-16 182-198
23	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 1.600e-23 126-157
24	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 1.600e-23 126-157
27	BL00239	Receptor tyrosine kinase class II proteins.	BL00239B 25.15 2.324e-16 91-139
28	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 3.250e-10 681-694 BL00018 7.41 6.400e-10 717-730
29	BL00018	EF-hand calcium-binding domain	BL00018 7.41 3.250e-10 681-694

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		proteins.	BL00018 7.41 6.400e-10 717-730
30	BL01113	C1q domain proteins.	BL01113A 17.99 9.308e-09 54-81
33	PD01168	SYNTHETASE LIGASE PROTEIN ALANYL.	PD01168L 9.47 1.667e-09 401-416
34	PD01168	SYNTHETASE LIGASE PROTEIN ALANYL.	PD01168L 9.47 1.667e-09 411-426
36	PR00426	C5A-ANAPHYLATOXIN RECEPTOR SIGNATURE	PR00426D 10.59 3.618e-12 110-122
37	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 2.049e-10 1080-1135
38	BL00350	MADS-box domain proteins.	BL00350 20.79 1.000e-40 1-55
40	BL00123	Alkaline phosphatase proteins.	BL00123B 19.31 1.000e-40 90-133 BL00123C 24.61 1.000e-40 145-195 BL00123E 22.25 1.000e-40 304-358 BL00123G 26.01 1.000e-40 438-488 BL00123F 19.03 8.714e-35 364-399 BL00123A 10.80 9.000e-24 52-77 BL00123D 12.73 1.000e-17 216-229
44	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.800e-14 346-359 PD00066 13.92 4.600e-14 486-499 PD00066 13.92 1.000e-13 374-387 PD00066 13.92 6.000e-13 458-471 PD00066 13.92 2.714e-12 234-247 PD00066 13.92 3.143e-12 430-443 PD00066 13.92 8.714e-12 514-527 PD00066 13.92 3.739e-11 402-415 PD00066 13.92 2.038e-10 318-331
45	DM00973	3 kw RESISTANCE BENOMYL YLL028W CYCLOHEXIMIDE.	DM00973A 21.17 2.946e-10 180-217
47	BL00649	G-protein coupled receptors family 2 proteins.	BL00649C 17.82 1.682e-10 475-501 BL00649B 20.68 7.387e-09 417-463
50	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-16 445-458 PD00066 13.92 5.846e-15 305-318 PD00066 13.92 1.000e-14 221-234 PD00066 13.92 1.000e-14 417-430 PD00066 13.92 2.800e-14 249-262 PD00066 13.92 2.800e-14 277-290 PD00066 13.92 8.800e-14 333-346 PD00066 13.92 9.400e-14 361-374 PD00066 13.92 4.000e-13 389-402 PD00066 13.92 6.571e-12 473-486
51	BL00226	Intermediate filaments proteins.	BL00226D 19.10 1.000e-40 417-464 BL00226B 23.86 3.348e-35 251-299 BL00226C 13.23 1.429e-24 316-347 BL00226A 12.77 1.857e-15 151-166
52	PR00217	43 KD POSTSYNAPTIC PROTEIN SIGNATURE	PR00217C 10.91 5.648e-09 133-149
53	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 1.000e-40 143-191 BL00232A 27.72 2.350e-28 49-82 BL00232B 32.79 7.052e-21 252-300 BL00232C 10.65 6.625e-20 250-268 BL00232B 32.79 1.314e-11 367-415 BL00232C 10.65 9.308e-10 470-488
54	BL00303	S-100/ICaBP type calcium binding	BL00303B 26.15 8.759e-23 125-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		protein.	162 BL00303A 21.77 1.000e-21 82-119
58	PR00378	INOSITOL PHOSPHATASE SIGNATURE	PR00378D 16.86 1.000e-15 242-261 PR00378B 13.80 9.250e-13 109-129
59	PR00425	BRADYKININ RECEPTOR SIGNATURE	PR00425C 13.23 9.040e-12 120-140
60	BL00280	Pancreatic trypsin inhibitor (Kunitz) family proteins.	BL00280 24.61 6.727e-38 238-282 BL00280 24.61 1.514e-30 294-338
65	BL01019	ADP-ribosylation factors family proteins.	BL01019A 13.20 1.222e-11 43-83
68	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237E 13.03 5.091e-13 188-212 PR00237G 19.63 7.207e-13 268-295 PR00237A 11.48 4.375e-11 24-49 PR00237C 15.69 3.057e-10 101-124 PR00237D 8.94 4.750e-10 137-159 PR00237F 13.57 5.364e-10 230-255 PR00237B 13.50 9.438e-10 57-79
70	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.938e-28 31-70
71	PR00830	ENDOPEPTIDASE LA (LON) SERINE PROTEASE (S16) SIGNATURE	PR00830A 8.41 8.759e-12 348-368
72	BL00120	Lipases, serine proteins.	BL00120B 11.37 2.149e-10 148-163
77	PR00753	1-AMINOCYCLOPROPANE-1-CARBOXYLATE SYNTHASE SIGNATURE	PR00753E 8.01 3.552e-11 191-216 PR00753D 6.85 2.778e-09 131-153
78	PR00506	D21 CLASS N6 ADENINE-SPECIFIC DNA METHYLTRANSFERASE SIGNATURE	PR00506C 19.40 8.017e-09 96-119
82	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.571e-16 436-467
84	BL00675	Sigma-54 interaction domain proteins ATP-binding region A proteins.	BL00675A 24.86 8.800e-10 256-300
85	BL00027	'Homeobox' domain proteins.	BL00027 26.43 2.286e-30 117-160
87	BL00250	TGF-beta family proteins.	BL00250A 21.24 6.786e-36 264-300 BL00250B 27.37 1.450e-26 328-364
91	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 9.250e-17 10-35 BL00215A 15.82 6.000e-16 221-246 BL00215A 15.82 7.857e-12 108-133 BL00215B 10.44 9.526e-11 168-181
92	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.526e-24 324-367
95	PR00094	ADENYLATE KINASE SIGNATURE	PR00094C 12.94 1.000e-08 119-136
96	PD02327	GLYCOPROTEIN ANTIGEN PRECURSOR IMMUNOGLO.	PD02327B 19.84 2.091e-09 143-165
97	BL00752	XPA protein.	BL00752B 19.17 7.309e-09 28-72
98	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876B 7.66 2.268e-10 135-149
99	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 9.824e-12 122-141
100	BL00027	'Homeobox' domain proteins.	BL00027 26.43 7.429e-31 118-161
101	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 6.870e-12 370-387 BL00028 16.07 6.885e-11 398-415 BL00028 16.07 8.269e-11 342-359 BL00028 16.07 4.300e-10 229-246

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00028 16.07 6.100e-10 258-275
102	PR00048	C2H2-TYPE ZINC FINGER SIGNATURE	PR00048A 10.52 7.750e-14 665-679 PR00048A 10.52 8.500e-14 581-595 PR00048A 10.52 9.250e-14 637-651 PR00048A 10.52 2.059e-12 609-623 PR00048A 10.52 2.588e-12 469-483 PR00048A 10.52 7.353e-12 553-567 PR00048A 10.52 2.895e-11 525-539 PR00048A 10.52 4.316e-11 441-455 PR00048A 10.52 5.263e-11 413-427 PR00048B 6.02 2.125e-10 569-579 PR00048B 6.02 4.938e-10 513-523 PR00048A 10.52 5.696e-10 497-511 PR00048B 6.02 8.875e-10 429-439 PR00048B 6.02 1.000e-09 457-467 PR00048B 6.02 6.684e-09 485-495
103	PR00195	DYNAMIN SIGNATURE	PR00195A 11.94 5.364e-22 31-50 PR00195B 9.47 1.783e-21 56-74 PR00195C 11.50 3.455e-21 126-144 PR00195D 11.76 8.714e-21 175-194 PR00195F 16.20 8.500e-20 217-237 PR00195E 9.82 8.650e-20 194-211
104	BL01113	C1q domain proteins.	BL01113A 17.99 1.865e-09 121-148 BL01113A 17.99 5.846e-09 82-109
105	BL00420	Speract receptor repeat proteins domain proteins.	BL00420A 20.42 6.400e-11 70-99 BL00420A 20.42 8.525e-10 73-102 BL00420A 20.42 5.708e-09 85-114
108	PR00860	VERTEBRATE METALLOTHIONEIN SIGNATURE	PR00860B 7.04 2.929e-20 27-41 PR00860A 5.46 5.500e-16 5-18 PR00860C 9.61 1.474e-14 41-51
112	BL01031	Heat shock hsp20 proteins family profile.	BL01031C 17.68 6.400e-10 122-147
114	DM01840	kw SPAC24B11.09 R07E5.13.	DM01840B 22.04 2.688e-40 59-103 DM01840A 10.95 9.571e-13 31-43
115	BL01126	Elongation factor Ts proteins.	BL01126A 18.48 2.317e-30 46-89 BL01126B 13.15 7.387e-19 116-135 BL01126C 9.20 9.735e-11 190-203
116	BL00216	Sugar transport proteins.	BL00216B 27.64 4.375e-21 35-85
118	BL00437	Catalase proximal heme-ligand proteins.	BL00437A 18.82 1.000e-40 49-101 BL00437B 16.28 1.000e-40 114-168 BL00437C 21.86 1.000e-40 190-239 BL00437D 25.72 1.000e-40 248-301 BL00437E 23.95 1.000e-40 327-379
119	BL00140	Ubiquitin carboxyl-terminal hydrolase family 1 cysteine activ.	BL00140D 22.64 8.274e-14 164-208 BL00140C 11.80 5.444e-10 77-102
120	BL00224	Clathrin light chain proteins.	BL00224B 16.94 6.712e-10 95-148
122	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 1.000e-40 16-62
123	PR00041	CAMP RESPONSE ELEMENT	PR00041D 7.95 2.906e-09 24-41

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		BINDING (CREB) PROTEIN SIGNATURE	
124	PR00041	CAMP RESPONSE ELEMENT BINDING (CREB) PROTEIN SIGNATURE	PR00041D 7.95 2.906e-09 24-41
125	BL00061	Short-chain dehydrogenases/reductases family proteins.	BL00061C 7.86 3.250e-10 212-222
126	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.400e-25 251-290
127	PR00318	ALPHA G-PROTEIN (TRANSDUCIN) SIGNATURE	PR00318D 16.28 1.900e-34 219-248 PR00318B 14.79 3.455e-27 168-191 PR00318C 12.09 7.000e-23 197-215 PR00318A 7.84 1.600e-19 35-51 PR00318E 7.23 2.500e-12 265-275
128	PR00927	ADENINE NUCLEOTIDE TRANSLOCATOR 1 SIGNATURE	PR00927E 14.93 9.743e-10 67-89 PR00927B 14.66 4.575e-09 69-91
130	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824B 9.21 7.750e-22 133-153
131	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824C 14.58 1.000e-40 166-204 BL00824D 14.04 1.621e-38 204-239 BL00824B 9.21 7.750e-22 133-153 BL00824E 12.49 1.000e-19 247-263
132	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 9.222e-13 1209-1228
133	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 9.222e-13 1168-1187
134	PR00708	ALPHA-1-ACID GLYCOPROTEIN SIGNATURE	PR00708D 14.67 1.000e-27 141-168 PR00708C 11.77 1.643e-25 98-120 PR00708B 15.15 2.174e-24 73-95 PR00708E 13.33 1.600e-21 189-207 PR00708A 14.40 2.636e-21 51-70
135	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 8.468e-13 126-145
136	PF00023	Ank repeat proteins.	PF00023A 16.03 3.250e-10 201-217
137	BL00471	Small cytokines (intercrine/chemokine) C-x-C subfamily signat.	BL00471 23.92 7.480e-10 42-90
140	PR00205	CADHERIN SIGNATURE	PR00205B 11.39 5.582e-10 328-346 PR00205B 11.39 9.018e-10 543-561
141	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 7.704e-09 976-1027
143	PR00979	TAFAZZIN SIGNATURE	PR00979E 10.83 5.950e-26 192-214 PR00979A 11.91 8.773e-25 63-83 PR00979C 12.16 6.400e-19 108-124 PR00979D 12.38 7.955e-19 170-185 PR00979F 10.14 3.382e-15 230-244 PR00979B 15.59 5.636e-15 94-106
145	DM00686	kw REPLICATION REP 28K 17.7K.	DM00686C 14.14 7.720e-09 111-131
146	PR00604	CLASS IA AND IB CYTOCHROME C SIGNATURE	PR00604D 15.86 1.000e-17 87-104 PR00604B 12.73 9.591e-16 57-73 PR00604C 10.21 8.200e-12 73-84 PR00604E 10.13 1.000e-11 106-117 PR00604A 11.13 8.800e-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			11 44-52 PR00604F 8.60 1.000e-10 123-132
147	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 3.864e-15 266-297 BL00107B 13.31 6.143e-11 335-351
148	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 8.448e-09 67-81
149	PR00069	ALDO-KETO REDUCTASE SIGNATURE	PR00069D 19.36 1.857e-30 187-217 PR00069A 16.01 7.429e-25 41-66 PR00069E 18.14 3.100e-22 235-260 PR00069C 16.03 7.000e-20 151-169 PR00069B 11.33 8.071e-19 101-120
150	BL00027	'Homeobox' domain proteins.	BL00027 26.43 2.688e-27 139-182
151	PD02906	SYNTHASE I PSEUDOURIDYLATE PSEUDOURIDINE LYASE TR.	PD02906C 24.17 7.070e-22 165-200 PD02906B 15.35 8.393e-15 114-127 PD02906A 10.84 6.500e-09 71-84
153	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479A 19.86 5.091e-12 891-914 BL00479B 12.57 1.837e-11 915-931
158	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.786e-31 143-186
160	BL00422	Granins proteins.	BL00422C 16.18 7.750e-12 420-448
162	PR00625	DNAJ PROTEIN FAMILY SIGNATURE	PR00625A 12.84 9.297e-11 62-82
164	BL01282	BIR repeat proteins.	BL01282B 30.49 6.182e-10 347-386
166	PR00860	VERTEBRATE METALLOTHIONEIN SIGNATURE	PR00860B 7.04 2.929e-20 83-97 PR00860A 5.46 1.000e-18 61-74 PR00860C 9.61 1.900e-15 97-107
167	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 7.052e-09 196-218
169	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 1.346e-39 316-353 BL00514G 15.98 2.241e-34 471-501 BL00514H 14.95 6.571e-27 510-535 BL00514E 14.28 1.273e-16 388-405 BL00514D 15.35 9.100e-15 369-382 BL00514B 16.42 4.857e-14 260-276 BL00514F 11.65 9.690e-14 416-431 BL00514A 11.68 8.200e-11 149-159
170	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 1.346e-39 268-305 BL00514G 15.98 2.241e-34 423-453 BL00514H 14.95 6.571e-27 462-487 BL00514E 14.28 1.273e-16 340-357 BL00514D 15.35 9.100e-15 321-334 BL00514B 16.42 4.857e-14 212-228 BL00514F 11.65 9.690e-14 368-383 BL00514A 11.68 8.200e-11 101-111
171	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514G 15.98 2.241e-34 385-415 BL00514H 14.95 6.571e-27 424-449 BL00514C 17.41 4.632e-24 230-267 BL00514E 14.28 1.273e-16 302-319 BL00514D 15.35 9.100e-15 283-296

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00514B 16.42 4.857e-14 212-228 BL00514F 11.65 9.690e-14 330-345 BL00514A 11.68 8.200e-11 101-111
173	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.400e-29 119-162
174	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 5.119e-15 1391-1404
176	BL00773	Chitinases family 19 proteins.	BL00773C 9.42 8.000e-09 2-16
182	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 9.163e-14 141-160
183	PD01937	DNA PROTEIN POLYMERASE ENDONUCLEASE DNA-.	PD01937A 6.68 3.475e-09 221-232
185	BL00845	CAP-Gly domain proteins.	BL00845 16.43 2.946e-23 247-272 BL00845 16.43 1.628e-21 107-132
186	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 6.538e-11 525-541
187	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 6.538e-11 497-513
188	DM01803	1 HERPESVIRUS GLYCOPROTEIN H.	DM01803A 10.51 1.000e-09 1081-1102
189	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 5.091e-15 69-82
190	PR00194	TROPOMYOSIN SIGNATURE	PR00194C 6.38 1.900e-35 145-174 PR00194E 8.74 3.250e-30 231-257 PR00194D 9.57 1.500e-26 175-199 PR00194B 10.24 5.200e-24 120-141 PR00194A 7.86 4.857e-21 84-102
192	PD02042	IRON-SULFUR ELECTRON TRANSPORT AROMATIC HYDROCARB.	PD02042B 16.75 5.154e-09 131-146 PD02042A 21.13 5.909e-09 94-121
193	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 2.200e-10 2-15
195	BL00463	Fungal Zn(2)-Cys(6) binuclear cluster domain proteins.	BL00463 8.22 5.071e-09 111-123
196	PR00118	BETA-LACTAMASE CLASS A SIGNATURE	PR00118F 16.42 9.386e-09 165-181
197	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 5.424e-09 234-267
198	BL00660	Band 4.1 family domain proteins.	BL00660A 31.50 5.500e-11 714-767
199	BL00282	Kazal serine protease inhibitors family proteins.	BL00282 16.88 8.820e-13 70-93
202	PR00009	TYPE I EGF SIGNATURE	PR00009A 14.15 5.345e-15 971-987 PR00009C 14.11 8.773e-13 996-1008 PR00009D 16.83 8.000e-11 1008-1018 PR00009C 14.11 1.882e-09 892-904
203	BL00025	P-type 'Trefoil' domain proteins.	BL00025 17.17 4.536e-19 38-59
205	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 7.300e-10 165-178
206	PR00168	SLOW VOLTAGE-GATED POTASSIUM CHANNEL SIGNATURE	PR00168D 12.88 6.865e-11 67-86
207	BL00025	P-type 'Trefoil' domain proteins.	BL00025 17.17 3.423e-20 39-60 BL00025 17.17 8.750e-16 88-109
209	BL00646	Ribosomal protein S13 proteins.	BL00646B 21.42 6.100e-30 110-143 BL00646A 25.82 6.192e-29 14-62
210	PR00138	MATRIXIN SIGNATURE	PR00138D 16.56 3.605e-25 279-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			305 PR00138C 16.41 3.000e-24 218-247 PR00138E 6.01 8.714e-13 314-328 PR00138A 15.14 9.538e-13 134-148 PR00138B 15.82 4.522e-12 188-204
211	DM01206	CORONAVIRUS NUCLEOCAPSID PROTEIN.	DM01206B 10.69 8.429e-12 386-406 DM01206B 10.69 1.247e-10 384-404 DM01206B 10.69 5.068e-10 388-408
212	PD01941	TRANSMEMBRANE COTRANSPORTER SYMP.	PD01941A 14.81 1.000e-40 163-217 PD01941B 15.02 9.705e-30 420-467 PD01941E 15.92 8.714e-23 837-884 PD01941C 19.96 8.200e-20 508-563 PD01941D 27.18 1.600e-16 661-710 PD01941F 28.52 9.645e-15 1005-1060
213	BL00362	Ribosomal protein S15 proteins.	BL00362 24.67 8.313e-09 330-373
214	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115Z 3.12 2.125e-09 1178-1227 BL00115Z 3.12 6.096e-09 1164-1213
215	BL00038	Myc-type, 'helix-loop-helix' dimerization domain proteins.	BL00038B 16.97 7.600e-18 125-146 BL00038A 13.61 1.474e-13 102-118
216	BL01108	Ribosomal protein L24 proteins.	BL01108A 20.33 2.241e-22 49-82 BL01108B 11.40 8.457e-10 96-107
217	PR00381	KINESIN LIGHT CHAIN SIGNATURE	PR00381A 9.55 1.321e-10 360-378
222	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 2.358e-26 1166-1203 BL00514G 15.98 9.000e-15 1289-1319 BL00514D 15.35 6.936e-12 1207-1220 BL00514F 11.65 4.288e-10 1253-1268 BL00514H 14.95 8.636e-10 1318-1343
223	BL00325	Actin-depolymerizing proteins.	BL00325B 21.66 1.000e-40 93-139 BL00325A 24.83 9.333e-24 61-93
224	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 1.450e-10 231-244
225	PF01329	Pterin 4 alpha carbinolamine dhydratase.	PF01329B 18.52 1.692e-18 67-92
228	BL00211	ABC transporters family proteins.	BL00211B 13.37 6.250e-18 1033-1065 BL00211B 13.37 8.875e-18 2045-2077 BL00211A 12.23 1.900e-09 931-943
230	PR00761	BINDIN PRECURSOR SIGNATURE	PR00761A 5.81 9.366e-09 275-292
231	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 3.500e-10 54-69
232	BL00412	Neuromodulin (GAP-43) proteins.	BL00412D 16.54 1.978e-10 109-160 BL00412D 16.54 4.122e-09 133-184
233	BL01210	Caveolins proteins.	BL01210B 13.92 8.129e-09 106-156
236	BL00939	Ribosomal protein L1e proteins.	BL00939F 17.27 5.393e-09 861-891
238	BL01252	Endogenous opioids neuropeptides precursors proteins.	BL01252D 18.25 3.571e-28 205-233 BL01252B 19.09 5.034e-27

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			37-67 BL01252C 18.10 1.621e-21 164-190 BL01252A 14.22 7.107e-18 14-34
239	BL00302	Eukaryotic initiation factor 5A hypusine proteins.	BL00302 14.81 1.000e-40 25-79
240	PR00420	AROMATIC-RING HYDROXYLASE (FLAVOPROTEIN MONOOXYGENASE) SIGNATURE	PR00420A 14.78 8.851e-13 26-49
241	PD02929	ADHESION GLYCOPROTEIN PRECURSOR I.	PD02929A 28.27 4.529e-09 235-289
243	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.527e-25 11-50
244	BL01270	Band 7 protein family proteins.	BL01270C 16.91 6.745e-17 115-144 BL01270B 18.74 6.857e-17 76-115 BL01270E 13.03 6.016e-15 182-211 BL01270D 20.87 9.160e-13 144-182
245	PF00791	Domain present in ZO-1 and Unc5-like netrin receptors.	PF00791B 28.49 6.305e-12 253-308 PF00791B 28.49 1.909e-11 427-482 PF00791B 28.49 2.651e-09 179-234 PF00791B 28.49 3.890e-09 112-167
246	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 2.500e-13 277-290 PD00066 13.92 9.143e-12 193-206 PD00066 13.92 5.304e-11 165-178 PD00066 13.92 6.478e-11 249-262 PD00066 13.92 3.423e-10 221-234
247	BL00406	Actins proteins.	BL00406D 12.58 6.400e-20 465-520 BL00406B 5.47 4.857e-14 249-304 BL00406E 8.44 1.000e-11 522-572 BL00406C 6.75 5.449e-11 313-368
248	BL00951	ER lumen protein retaining receptor proteins.	BL00951C 19.35 1.000e-40 112-161 BL00951A 15.10 7.750e-39 21-57 BL00951D 13.94 6.000e-38 161-196 BL00951B 14.23 3.100e-31 57-88
252	BL01113	C1q domain proteins.	BL01113A 17.99 9.129e-15 200-227 BL01113A 17.99 4.818e-14 194-221 BL01113A 17.99 7.818e-14 182-209 BL01113A 17.99 1.730e-13 185-212 BL01113A 17.99 6.595e-13 191-218 BL01113A 17.99 6.077e-12 203-230 BL01113A 17.99 9.182e-11 179-206 BL01113A 17.99 2.532e-10 176-203 BL01113A 17.99 9.043e-10 218-245 BL01113A 17.99 9.426e-10 209-236 BL01113A 17.99 4.115e-09 137-164
257	BL00845	CAP-Gly domain proteins.	BL00845 16.43 1.837e-21 466-491
259	PR00248	METABOTROPIC GLUTAMATE GPCR SIGNATURE	PR00248G 12.67 2.688e-09 53-78
260	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 3.400e-10 441-452 BL00678 9.67 5.800e-10 481-492 BL00678 9.67 8.800e-10 358-369
261	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 3.400e-10 415-426 BL00678 9.67 5.800e-10 455-466

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00678 9.67 8.800e-10 332-343
262	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 3.400e-10 468-479 BL00678 9.67 5.800e-10 508-519 BL00678 9.67 8.800e-10 385-396
263	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002B 15.18 2.200e-10 415-429
264	BL00049	Ribosomal protein L14 proteins.	BL00049C 17.38 3.040e-12 94-130
265	PD01469	GLYCOPROTEIN PROTEIN PRECURSOR SA.	PD01469 20.59 2.091e-14 438-470
266	PD01469	GLYCOPROTEIN PROTEIN PRECURSOR SA.	PD01469 20.59 2.091e-14 279-311
267	BL00567	Phosphoribulokinase proteins.	BL00567A 10.66 1.161e-12 36-55
269	BL00049	Ribosomal protein L14 proteins.	BL00049C 17.38 2.688e-28 92-128 BL00049B 18.42 6.806e-24 54-86 BL00049A 13.86 8.333e-19 19-42 BL00049D 13.47 5.765e-12 129-140
272	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 9.735e-12 14-58
273	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 1.911e-09 819-832
275	PR00179	LIPOCALIN SIGNATURE	PR00179B 9.56 2.895e-13 124-137 PR00179A 13.78 3.250e-11 36-49 PR00179C 19.02 6.040e-11 154-170
276	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 8.364e-17 22-44 PR00449C 17.27 1.000e-13 62-85 PR00449E 13.50 4.000e-12 172-195 PR00449B 14.34 5.680e-10 45-62
277	BL00140	Ubiquitin carboxyl-terminal hydrolase family 1 cysteine activ.	BL00140D 22.64 1.000e-40 161-205 BL00140C 11.80 9.053e-30 79-104 BL00140A 15.96 9.400e-28 5-35 BL00140B 12.29 4.649e-17 37-55
278	PD02712	ELEMENT TRANSPOSASE FOR TRANSPOSON TRANSPOSABLE.	PD02712A 23.03 8.013e-09 47-83
279	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 1.474e-09 100-111
282	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 4.767e-21 864-898
283	BL00048	Protamine P1 proteins.	BL00048 6.39 9.550e-09 56-83
286	PR00081	GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE	PR00081A 10.53 1.878e-11 36-54
287	PR00310	ANTI-PROLIFERATIVE PROTEIN BTG1 FAMILY SIGNATURE	PR00310B 10.59 4.231e-17 29-59 PR00310D 9.10 6.679e-16 89-119
289	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.000e-36 37-76
293	BL00979	G-protein coupled receptors family 3 proteins.	BL00979L 20.63 3.800e-12 111-152
295	PD02411	PROTEIN TRANSCRIPTION REGULATION NUCLEAR.	PD02411 21.89 7.000e-16 195-229
296	BL01064	Pyridoxamine 5'-phosphate oxidase proteins.	BL01064A 27.84 8.313e-28 77-129 BL01064C 15.22 7.136e-25 202-235
297	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030A 14.39 2.929e-13 37-56 BL00030B 7.03 1.900e-11 167-177 BL00030A 14.39 2.000e-10 128-147

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
298	BL01183	ubiE/COQ5 methyltransferase family proteins.	BL01183B 21.31 6.660e-12 143-188
299	BL01279	Protein-L-isoaspartate(D-aspartate) O-methyltransferase signa.	BL01279A 24.27 5.862e-11 57-105
301	BL00191	Cytochrome b5 family, heme-binding domain proteins.	BL00191K 17.38 4.951e-27 184-228 BL00191J 11.37 6.447e-17 128-150
302	DM00892	3 RETROVIRAL PROTEINASE.	DM00892C 23.55 3.893e-16 33-67
306	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 2.988e-09 416-451
307	PR00245	OLFACTORY RECEPTOR SIGNATURE	PR00245A 18.03 4.818e-21 59-81 PR00245C 7.84 5.154e-20 238-254 PR00245D 10.47 4.000e-15 274-286 PR00245B 10.38 8.200e-15 177-192 PR00245E 12.40 5.714e-12 291-306
309	BL00203	Vertebrate metallothioneins proteins.	BL00203 13.94 2.245e-10 612-658
310	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 7.632e-23 119-159 BL00237C 13.19 3.864e-15 251-278 BL00237D 11.23 3.739e-12 312-329
311	BL00380	Rhodanese proteins.	BL00380D 15.90 8.200e-28 110-136 BL00380G 11.26 5.800e-16 267-280 BL00380B 14.77 7.000e-14 49-62 BL00380F 9.76 5.886e-13 203-214 BL00380C 15.67 7.387e-13 82-98 BL00380E 12.44 7.000e-11 181-193 BL00380A 10.48 1.000e-09 10-20
312	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227B 19.29 1.000e-40 50-105 BL00227C 25.48 1.000e-40 111-163 BL00227D 18.46 1.000e-40 220-274 BL00227F 21.16 1.000e-40 372-426 BL00227A 24.55 3.250e-39 1-35 BL00227E 24.15 8.500e-34 324-359
327	BL00232	Cadherins extracellular repeat proteins domain proteins.	BL00232B 32.79 7.362e-21 225-273 BL00232B 32.79 2.588e-17 435-483 BL00232B 32.79 6.301e-15 116-164 BL00232B 32.79 6.769e-13 330-378 BL00232C 10.65 9.341e-12 223-241 BL00232C 10.65 5.696e-11 328-346 BL00232C 10.65 3.942e-10 433-451
329	PD02749	TRANSCRIPTION PROTEIN FACTOR BTF3 REGULATION NUCL.	PD02749B 12.75 2.241e-37 35-71 PD02749C 13.96 4.892e-28 87-121 PD02749A 9.56 6.000e-15 2-15
330	PR00391	PHOSPHATIDYLINOSITOL TRANSFER PROTEIN SIGNATURE	PR00391E 12.50 7.785e-15 211-231 PR00391B 8.39 1.000e-13 83-104 PR00391D 12.21 9.328e-13 191-207 PR00391A 7.83 5.390e-11 16-36
332	BL01030	RNA polymerases M / 15 Kd subunits proteins.	BL01030 23.44 1.818e-23 87-125
337	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.929e-32 6-45
340	PD02711	SYNTHASE	PD02711B 14.26 1.973e-20 944-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		PHOSPHORIBOSYLFORMYLGLY.	968
343	BL00223	Annexins repeat proteins domain proteins.	BL00223C 24.79 1.000e-40 245-300 BL00223B 28.47 8.714e-38 168-218 BL00223A 15.59 8.250e-27 98-132 BL00223A 15.59 8.750e-27 26-60 BL00223C 24.79 9.438e-16 13-68 BL00223C 24.79 2.735e-15 85-140 BL00223A 15.59 2.253e-11 258-292
346	PR00345	STATHMIN FAMILY SIGNATURE	PR00345B 7.12 2.800e-28 81-110 PR00345E 8.54 7.652e-28 158-183 PR00345C 4.54 9.100e-28 110-134 PR00345D 10.97 1.964e-24 134-158 PR00345A 13.46 5.645e-16 52-71
347	BL00586	Ribosomal protein L16 proteins.	BL00586B 17.00 3.215e-15 184-221
348	PR00388	3',5'-CYCLIC NUCLEOTIDE CLASS II PHOSPHODIESTERASE SIGNATURE	PR00388A 10.45 2.778e-09 86-105
351	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 3.118e-11 160-173 BL00018 7.41 2.350e-10 244-257
354	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 1.947e-09 256-267
358	DM01206	CORONAVIRUS NUCLEOCAPSID PROTEIN.	DM01206B 10.69 3.278e-09 175-195 DM01206B 10.69 6.696e-09 183-203 DM01206B 10.69 8.633e-09 132-152 DM01206B 10.69 8.861e-09 181-201 DM01206B 10.69 9.316e-09 177-197
361	PD01498	OXIDASE BIOSYNTHESIS OXIDOREDUCTASE PORP.	PD01498C 24.90 6.880e-14 219-263
362	PD01498	OXIDASE BIOSYNTHESIS OXIDOREDUCTASE PORP.	PD01498C 24.90 6.880e-14 219-263
365	BL00178	Aminoacyl-transfer RNA synthetases class-I proteins.	BL00178B 7.11 1.000e-11 589-600 BL00178A 14.23 8.500e-09 46-56
366	BL00523	Sulfatases proteins.	BL00523E 19.27 1.000e-23 318-348 BL00523A 13.36 5.500e-16 30-47 BL00523B 8.64 1.964e-13 78-90 BL00523C 12.64 9.625e-13 129-140 BL00523G 9.46 5.500e-10 506-516
369	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.818e-09 21-52
370	BL00880	Acyl-CoA-binding protein.	BL00880 17.52 1.000e-40 75-125
371	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 1.000e-23 276-307 BL00107B 13.31 1.692e-12 342-358
372	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 6.602e-11 326-347 PR00211B 0.86 6.106e-10 320-341 PR00211B 0.86 3.167e-09 333-354
373	BL00279	Membrane attack complex components / perforin proteins.	BL00279E 37.11 9.349e-10 749-797
375	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 1.231e-33 10-49
377	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.563e-28 10-49
379	BL00598	Chromo domain proteins.	BL00598 14.45 5.781e-16 3-25

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
380	PR00413	HALOACID DEHALOGENASE/EPOXIDE HYDROLASE FAMILY SIGNATURE	PR00413D 11.28 8.941e-09 864-878
383	PR00413	HALOACID DEHALOGENASE/EPOXIDE HYDROLASE FAMILY SIGNATURE	PR00413D 11.28 8.941e-09 864-878
387	BL01060	Flagella transport protein fliP family proteins.	BL01060A 15.65 1.535e-09 131-174
388	PR00209	ALPHA/BETA GLIADIN FAMILY SIGNATURE	PR00209B 4.88 6.318e-11 1009-1028
389	PR00837	ALLERGEN V5/TPX-1 FAMILY SIGNATURE	PR00837B 11.64 1.000e-10 469-483
391	BL00240	Receptor tyrosine kinase class III proteins.	BL00240B 24.70 7.907e-10 118-142
392	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014D 12.04 8.412e-10 691-706
393	PR00014	FIBRONECTIN TYPE III REPEAT SIGNATURE	PR00014D 12.04 8.412e-10 706-721
394	BL01209	LDL-receptor class A (LDLRA) domain proteins.	BL01209 9.31 3.368e-15 47-60 BL01209 9.31 5.500e-13 92-105
395	BL00634	Ribosomal protein L30 proteins.	BL00634 34.38 4.090e-13 70-121
396	BL01013	Oxysterol-binding protein family proteins.	BL01013D 26.81 8.000e-26 358-402 BL01013A 25.14 7.231e-21 45-81 BL01013C 9.97 1.000e-13 132-142 BL01013B 11.33 1.000e-11 110-121
397	BL00930	Peripherin / rom-1 proteins.	BL00930E 17.80 1.000e-40 56-92 BL00930D 9.12 4.632e-37 12-56 BL00930F 16.91 2.800e-36 92-133
400	PR00780	LEUSERPIN 2 SIGNATURE	PR00780B 4.89 4.491e-09 262-285
401	PR00819	CBXX/CFQX SUPERFAMILY SIGNATURE	PR00819B 10.83 7.158e-11 4-20
403	BL00381	Endopeptidase Clp serine proteins.	BL00381C 23.84 1.250e-32 150-194 BL00381A 16.48 2.286e-22 74-111 BL00381B 21.42 8.326e-14 78-130
405	BL01105	Ribosomal protein L35Ae proteins.	BL01105A 17.37 1.000e-40 4-49 BL01105B 12.95 1.000e-40 68-108
406	BL00344	GATA-type zinc finger domain proteins.	BL00344 17.99 7.000e-12 814-852
407	PR00211	GLUTELIN SIGNATURE	PR00211B 0.86 9.750e-09 73-94
409	PR00910	LUTEOVIRUS ORF6 PROTEIN SIGNATURE	PR00910A 2.51 4.321e-09 9-22
410	BL00762	WHEP-TRS domain proteins.	BL00762A 23.43 1.000e-28 752-789 BL00762A 23.43 4.400e-21 903-940 BL00762A 23.43 5.415e-18 825-862 BL00762B 16.14 8.759e-12 1154-1168
412	BL00690	DEAH-box subfamily ATP-dependent helicases proteins.	BL00690B 13.38 5.320e-15 262-280 BL00690A 6.87 1.818e-13 230-240
415	BL00227	Tubulin subunits alpha, beta, and gamma proteins.	BL00227B 19.29 1.000e-40 52-107 BL00227C 25.48 1.000e-40 113-165 BL00227D 18.46 1.000e-40 222-276 BL00227F 21.16 1.000e-40 382-436 BL00227E 24.15 1.750e-34 326-361

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00227A 24.55 1.000e-33 1-35
416	PF00992	Troponin.	PF00992A 16.67 1.711e-09 557-592
418	BL00541	Nuclear transition protein 1 proteins.	BL00541 8.44 9.875e-09 256-310
419	BL00541	Nuclear transition protein 1 proteins.	BL00541 8.44 9.875e-09 197-251
420	PF00856	SET domain proteins.	PF00856A 26.14 9.074e-13 901-938 PF00856B 16.42 2.397e-12 951-973
421	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 8.200e-12 33-44
423	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.600e-30 130-169
424	PF00564	Octicosapeptide repeat proteins.	PF00564B 24.74 1.305e-17 421-472
426	PR00988	URIDINE KINASE SIGNATURE	PR00988A 6.39 4.569e-12 3-21
427	PR00988	URIDINE KINASE SIGNATURE	PR00988A 6.39 4.569e-12 3-21
428	BL00478	LIM domain proteins.	BL00478B 14.79 3.250e-13 115-130 BL00478B 14.79 9.036e-13 50-65
431	BL00282	Kazal serine protease inhibitors family proteins.	BL00282 16.88 8.875e-12 464-487
432	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 7.800e-18 316-357 PD00930A 25.62 9.617e-12 125-151 PD00930B 33.72 2.521e-10 214-255
433	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.649e-34 34-73
434	PR00449	TRANSFORMING PROTEIN P21 RAS SIGNATURE	PR00449A 13.20 7.563e-11 56-78
436	PR00120	H+-TRANSPORTING ATPASE (PROTON PUMP) SIGNATURE	PR00120C 9.90 5.800e-19 705-722
437	BL00115	Eukaryotic RNA polymerase II heptapeptide repeat proteins.	BL00115T 8.45 7.273e-29 1208-1242 BL00115Q 18.08 2.776e-21 953-983 BL00115Y 11.86 8.000e-17 1604-1650 BL00115M 19.19 8.130e-16 731-774 BL00115H 14.34 9.392e-16 463-496 BL00115A 15.44 7.414e-15 43-82 BL00115R 6.50 6.128e-14 983-1010 BL00115J 16.71 9.289e-14 591-617 BL00115I 8.33 4.336e-13 535-590 BL00115L 12.25 5.939e-13 662-694 BL00115G 11.65 6.011e-13 435-463 BL00115K 15.03 3.417e-10 617-659 BL00115O 16.76 5.805e-10 863-913 BL00115P 11.54 7.538e-10 913-953 BL00115S 18.24 7.968e-10 1010-1052 BL00115U 10.34 4.475e-09 1242-1265
438	PF00628	PHD-finger.	PF00628 15.84 4.536e-10 219-234
440	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.351e-34 10-49
441	PR00309	ARRESTIN SIGNATURE	PR00309A 9.68 5.250e-24 32-55 PR00309D 7.09 4.938e-23 290-309 PR00309B 7.81 2.800e-21 69-88 PR00309C 8.22 1.621e-19 165-183 PR00309E 9.82 9.438e-15 374-389
442	BL00600	Aminotransferases class-III pyridoxal-	BL00600B 19.60 7.324e-14 103-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		phosphate attachment si.	129 BL00600G 12.43 2.125e-12 306-325 BL00600F 8.77 8.105e-12 271-284 BL00600E 16.43 3.167e-11 228-257 BL00600D 8.71 8.650e-09 207-221
443	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 3.160e-18 69-87
444	BL00349	CTF/NF-I proteins.	BL00349A 10.07 1.000e-40 8-54 BL00349C 9.33 1.000e-40 82-125 BL00349E 10.79 1.000e-40 152-195 BL00349F 11.81 1.000e-40 213-255 BL00349H 15.70 7.387e-36 361-399 BL00349B 10.51 2.227e-34 54-82 BL00349D 11.70 9.100e-34 125-152 BL00349G 19.72 5.781e-30 323-356
445	BL00154	E1-E2 ATPases phosphorylation site proteins.	BL00154F 8.23 8.941e-21 271-295 BL00154E 20.37 2.620e-15 124-165
448	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 4.882e-11 82-115 DM00215 19.43 6.492e-09 87-120
451	BL01283	T-box domain proteins.	BL01283A 24.15 3.100e-40 112-160 BL01283D 11.70 6.000e-39 253-286 BL01283B 23.17 6.538e-38 170-212 BL01283C 13.05 7.750e-19 222-236
452	PR00420	AROMATIC-RING HYDROXYLASE (FLAVOPROTEIN MONOOXYGENASE) SIGNATURE	PR00420A 14.78 2.579e-11 3-26
453	PR00162	RIESKE 2FE-2S SUBUNIT SIGNATURE	PR00162B 12.77 7.429e-17 215-228 PR00162A 9.35 2.324e-14 193-205 PR00162C 8.10 7.120e-14 227-240
454	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.000e-30 87-126
456	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.333e-18 1149-1192
457	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.737e-24 16-55
459	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290A 20.89 1.529e-14 154-177 BL00290B 13.17 9.000e-12 214-232
460	PR00413	HALOACID DEHALOGENASE/EPOXIDE HYDROLASE FAMILY SIGNATURE	PR00413F 14.91 7.333e-11 193-214 PR00413E 15.78 5.714e-09 175-192
463	PR00759	BASIC PROTEASE (KUNITZ-TYPE) INHIBITOR FAMILY SIGNATURE	PR00759B 11.26 8.385e-09 74-85
466	BL00019	Actinin-type actin-binding domain proteins.	BL00019D 15.33 4.200e-19 300-330
467	BL00019	Actinin-type actin-binding domain proteins.	BL00019D 15.33 4.200e-19 300-330
469	PR00153	CYCLOPHILIN PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SIGNATURE	PR00153D 11.99 3.250e-15 510-523 PR00153C 11.01 4.682e-14 495-511 PR00153E 9.10 8.548e-14 523-539 PR00153B 11.57 1.720e-13 452-465
470	BL00491	Aminopeptidase P and proline dipeptidase proteins.	BL00491C 12.15 3.912e-09 557-572
471	PD00289	PROTEIN SH3 DOMAIN REPEAT	PD00289 9.97 1.000e-14 1482-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		PRESYNA.	1496 PD00289 9.97 8.650e-11 1122-1136
474	BL50040	Elongation factor 1 gamma chain profile.	BL50040D 17.41 1.000e-40 279-329 BL50040E 18.79 1.000e-40 333-388 BL50040F 18.99 5.320e-40 390-428 BL50040C 22.62 3.739e-38 141-184 BL50040B 13.65 7.000e-30 59-85 BL50040A 12.98 1.450e-14 10-22
475	BL01144	Ribosomal protein L31e proteins.	BL01144 25.07 1.000e-40 22-74
476	PR00007	COMPLEMENT C1Q DOMAIN SIGNATURE	PR00007C 15.60 2.421e-21 589-611 PR00007B 14.16 3.500e-21 544-564 PR00007A 19.33 6.897e-20 517-544 PR00007D 9.64 6.571e-12 623-634
477	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 5.846e-10 170-189
479	DM01970	0 kw ZK632.12 YDR313C ENDOSOMAL III.	DM01970B 8.60 9.500e-17 967-980
480	PR00868	DNA-POLYMERASE FAMILY A (POL I) SIGNATURE	PR00868C 13.76 5.688e-17 284-308 PR00868A 16.33 3.186e-13 224-247 PR00868H 12.51 3.388e-13 431-448 PR00868I 10.87 7.938e-11 462-476 PR00868E 13.19 1.608e-10 340-366
481	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.182e-22 53-96
482	BL00061	Short-chain dehydrogenases/reductases family proteins.	BL00061B 25.79 3.647e-21 188-226
483	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 1.750e-12 1032-1051
485	PF00023	Ank repeat proteins.	PF00023A 16.03 9.625e-10 760-776 PF00023A 16.03 3.571e-09 715-731
486	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870B 18.83 9.262e-20 103-136 PD02870D 15.74 9.426e-09 201-236
487	PR00370	FLAVIN-CONTAINING MONOOXYGENASE (FMO) SIGNATURE	PR00370G 10.45 3.769e-28 471-493 PR00370B 10.91 1.000e-24 27-46 PR00370C 12.72 4.000e-21 140-157 PR00370E 11.96 9.229e-21 320-339 PR00370D 16.33 1.750e-20 185-204 PR00370F 17.75 7.395e-20 375-395 PR00370A 3.35 2.038e-18 4-20
489	PD01675	GLYCOPROTEIN MAJOR ENVELOPE PROBABLE U3.	PD01675C 19.89 2.330e-10 55-89
492	BL00211	ABC transporters family proteins.	BL00211A 12.23 5.050e-09 45-57
493	BL00211	ABC transporters family proteins.	BL00211A 12.23 5.050e-09 45-57
494	BL00211	ABC transporters family proteins.	BL00211A 12.23 5.050e-09 58-70
495	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.786e-12 509-552 BL00027 26.43 9.143e-12 319-362 BL00027 26.43 2.600e-11 627-670 BL00027 26.43 3.625e-10 779-822
497	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.800e-22 214-245 BL00107B 13.31 1.000e-13 281-297 BL00107A 18.39 3.520e-13 583-614 BL00107B 13.31 8.615e-12 652-668
499	BL00383	Tyrosine specific protein phosphatases	BL00383E 10.35 1.000e-14 1902-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		proteins.	1913 BL00383D 11.92 3.077e-14 1862-1875 BL00383A 13.34 5.500e-14 1730-1745 BL00383C 10.10 2.000e-13 1785-1796 BL00383F 15.51 9.069e-12 1940-1956 BL00383B 7.61 1.692e-11 1755-1764
501	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 1.360e-09 136-150 PR00019A 11.19 1.667e-09 91-105 PR00019B 11.36 4.600e-09 160-174
503	BL00226	Intermediate filaments proteins.	BL00226D 19.10 1.000e-40 367-414 BL00226B 23.86 6.143e-27 195-243 BL00226A 12.77 7.840e-14 96-111 BL00226C 13.23 2.600e-13 309-340 BL00226C 13.23 6.143e-12 266-297 BL00226B 23.86 1.209e-09 146-194
505	PD02407	3-BISPHOSPHOGLYCERATE-INDEPENDENT PHOSPHOGLYCER.	PD02407F 7.61 6.739e-09 916-930
506	PF00632	HECT-domain (ubiquitin-transferase).	PF00632C 20.66 9.830e-19 991-1023 PF00632B 18.45 1.155e-11 940-968
507	BL01082	Ribosomal protein L7Ae proteins.	BL01082 20.37 4.273e-20 76-116
508	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 2.421e-09 493-504
509	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 2.421e-09 473-484
510	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320B 12.19 4.774e-11 567-582 PR00320B 12.19 5.886e-10 763-778 PR00320C 13.01 6.760e-10 567-582 PR00320A 16.74 7.618e-10 846-861 PR00320A 16.74 3.415e-09 763-778 PR00320A 16.74 6.268e-09 567-582
511	BL00479	Phorbol esters / diacylglycerol binding domain proteins.	BL00479C 12.01 3.250e-12 170-183
512	BL50058	G-protein gamma subunit profile.	BL50058 27.23 7.494e-09 10-58
513	BL00524	Somatomedin B domain proteins.	BL00524A 9.65 8.925e-14 80-101
515	BL00041	Bacterial regulatory proteins, araC family proteins.	BL00041 23.99 1.964e-19 492-524
516	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.500e-13 391-404
517	BL00415	Synapsins proteins.	BL00415E 4.82 9.291e-09 959-996
518	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 9.471e-12 126-145
519	BL00290	Immunoglobulins and major histocompatibility complex proteins.	BL00290B 13.17 4.750e-09 47-65
522	PR00505	D12 CLASS N6 ADENINE-SPECIFIC DNA METHYLTRANSFERASE SIGNATURE	PR00505A 14.15 7.128e-09 364-381
525	BL00312	Glycophorin A proteins.	BL00312B 9.22 5.781e-10 891-920
528	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.500e-32 16-55
529	PR00254	NICOTINIC ACETYLCHOLINE RECEPTOR SIGNATURE	PR00254D 15.50 4.000e-17 131-150 PR00254A 11.23 4.706e-14 61-78 PR00254C 11.36 4.000e-12

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			113-126 PR00254B 12.97 1.486e-11 95-110
531	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 6.870e-16 787-810
532	PR00193	MYOSIN HEAVY CHAIN SIGNATURE	PR00193D 14.36 3.143e-34 447-476 PR00193C 12.60 7.632e-32 216-244 PR00193B 11.69 7.750e-29 167-193 PR00193A 15.41 2.588e-22 111-131 PR00193E 19.47 2.200e-21 501-530
533	PD02870	RECEPTOR INTERLEUKIN-1 PRECURSOR.	PD02870B 18.83 5.596e-09 348-381
535	PR00683	SPECTRIN PLECKSTRIN HOMOLOGY DOMAIN SIGNATURE	PR00683D 15.87 2.452e-10 465-484
536	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.684e-24 164-207
538	PR00239	MOLLUSCAN RHODOPSIN C-TERMINAL TAIL SIGNATURE	PR00239E 1.58 2.739e-09 225-237
539	BL00406	Actins proteins.	BL00406C 6.75 1.000e-40 157-212 BL00406B 5.47 6.143e-37 90-145 BL00406D 12.58 4.600e-36 291-346 BL00406E 8.44 2.200e-33 364-414 BL00406A 9.95 4.441e-23 7-42
540	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456E 3.06 9.625e-10 44-59
541	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456E 3.06 9.625e-10 44-59
542	PF00023	Ank repeat proteins.	PF00023A 16.03 7.857e-11 138-154
544	PF00642	Zinc finger C-x8-C-x5-C-x3-H type (and similar).	PF00642 11.59 9.082e-10 838-849
546	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 4.115e-10 104-115
547	BL01226	Hydroxymethylglutaryl-coenzyme A synthase proteins.	BL01226A 13.79 1.000e-40 50-89 BL01226C 13.51 1.000e-40 127-167 BL01226D 11.60 1.000e-40 174-210 BL01226E 13.74 1.000e-40 212-253 BL01226H 17.74 1.000e-40 386-434 BL01226I 25.06 1.000e-40 460-508 BL01226G 15.76 3.483e-32 292-321 BL01226B 13.35 1.818e-31 95-127 BL01226F 9.78 8.714e-23 253-271
549	BL00964	Syndecans proteins.	BL00964B 12.05 2.426e-10 1246-1289
551	DM01930	2 kw FINGER SMCX SMCY YDR096W.	DM01930E 15.41 1.367e-37 170-215 DM01930F 14.16 8.232e-28 267-303 DM01930B 19.86 9.163e-10 37-71
552	BL00195	Glutaredoxin proteins.	BL00195B 15.31 7.158e-09 9-29
554	BL00383	Tyrosine specific protein phosphatases proteins.	BL00383E 10.35 2.756e-12 436-447
555	PR00403	WW DOMAIN SIGNATURE	PR00403B 12.19 7.612e-11 122-137 PR00403A 16.82 3.912e-10 107-121 PR00403B 12.19 2.068e-09 76-91
558	PR00380	KINESIN HEAVY CHAIN SIGNATURE	PR00380A 14.18 2.714e-26 76-98 PR00380D 9.93 3.000e-24 275-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			297 PR00380C 13.18 5.154e-20 226-245 PR00380B 12.64 9.400e-20 195-213
559	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 5.333e-09 522-531
561	PD01795	PROTEIN AMINOPEPTIDASE PRECURSOR HYDROLASE SIGNA.	PD01795B 11.56 2.333e-12 159-172 PD01795A 10.27 1.000e-09 135-144
562	PD01795	PROTEIN AMINOPEPTIDASE PRECURSOR HYDROLASE SIGNA.	PD01795B 11.56 2.333e-12 110-123 PD01795A 10.27 1.000e-09 86-95
563	BL00018	EF-hand calcium-binding domain proteins.	BL00018 7.41 1.391e-09 41-54
565	BL00348	p53 tumor antigen proteins.	BL00348F 23.19 4.143e-09 188-231
567	PD00301	PROTEIN REPEAT MUSCLE CALCIUM-BI.	PD00301B 5.49 4.115e-09 284-295
569	PF00850	Histone deacetylase family.	PF00850E 8.88 6.553e-21 756-782 PF00850D 14.76 1.519e-16 722-746 PF00850F 15.70 1.118e-11 794-827 PF00850G 22.75 8.375e-11 833-875
570	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 4.960e-10 137-151
571	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 8.800e-11 44-53
573	BL00299	Ubiquitin domain proteins.	BL00299 28.84 1.123e-11 123-175
574	PF01140	Matrix protein (MA), p15.	PF01140D 15.54 3.700e-10 986-1021
576	BL00284	Serpins proteins.	BL00284C 28.56 5.200e-26 200-242 BL00284A 15.64 4.913e-18 71-95 BL00284B 17.99 7.261e-15 173-194 BL00284D 16.34 5.846e-13 306-333 BL00284E 19.15 7.429e-12 387-412
579	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.553e-29 15-54
580	BL50001	Src homology 2 (SH2) domain proteins profile.	BL50001B 17.40 4.500e-12 1010-1031
581	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 3.189e-22 608-649 PD00930A 25.62 6,806e-17 505-531
584	BL00612	Osteonectin domain proteins.	BL00612B 11.35 2.034e-11 93-126
585	DM01551	kw OSTEOINDUCTIVE YOPM MEMBRANE OUTER.	DM01551C 14.62 8.859e-10 102-122
586	PF00628	PHD-finger.	PF00628 15.84 3.455e-12 235-250
587	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.063e-10 85-128
588	PR00326	GTP1/OBG GTP-BINDING PROTEIN FAMILY SIGNATURE	PR00326A 8.75 7.525e-16 227-248 PR00326C 9.79 6.760e-15 276-292 PR00326D 19.09 6.657e-13 293-312 PR00326B 16.74 9.229e-13 248-267
589	BL00422	Granins proteins.	BL00422A 28.34 7.429e-09 2349-2378
590	BL00415	Synapsins proteins.	BL00415N 4.29 9.794e-10 295-339
591	BL00128	Alpha-lactalbumin / lysozyme C proteins.	BL00128A 20.76 3.423e-13 35-65 BL00128C 19.34 2.980e-11 110-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			132
596	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 3.136e-09 31-46
597	DM00547	1 kw CHROMO BROMODOMAIN SHADOW GLOBAL.	DM00547C 17.30 1.667e-19 207-229 DM00547E 13.94 6.200e-18 319-342 DM00547B 11.28 1.000e-17 179-193 DM00547D 11.60 9.250e-13 289-303 DM00547F 23.43 6.727e-12 679-726 DM00547A 12.38 4.818e-11 158-170
600	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 1.882e-27 13-52
601	BL00192	Cytochrome b/b6 heme-ligand proteins.	BL00192A 11.90 6.400e-09 390-430
602	BL00936	Ribosomal protein L35 proteins.	BL00936B 27.27 8.615e-09 118-157
603	BL00936	Ribosomal protein L35 proteins.	BL00936B 27.27 8.615e-09 118-157
606	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 7.300e-10 292-306 PR00019A 11.19 5.667e-09 323-337
607	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 7.300e-10 292-306 PR00019A 11.19 5.667e-09 323-337
608	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 9.500e-12 168-183 PR00320A 16.74 2.853e-10 60-75 PR00320A 16.74 4.706e-10 14-29 PR00320C 13.01 5.320e-10 60-75 PR00320C 13.01 5.680e-10 14-29 PR00320A 16.74 6.049e-09 217-232 PR00320B 12.19 8.875e-09 168-183
610	BL00750	Chaperonins TCP-1 proteins.	BL00750B 16.17 1.000e-40 70-120 BL00750A 20.07 6.211e-37 26-69 BL00750G 20.12 8.800e-31 431-471 BL00750F 18.40 5.125e-30 370-411 BL00750E 24.59 8.650e-29 295-332 BL00750H 21.44 1.000e-27 489-524 BL00750C 25.65 5.345e-17 149-181 BL00750D 16.16 6.318e-14 203-222
613	BL00766	Tetrahydrofolate dehydrogenase/cyclohydrolase proteins.	BL00766B 24.49 1.000e-40 142-190 BL00766E 13.78 1.000e-40 322-359 BL00766C 25.86 5.500e-39 208-256 BL00766D 17.05 4.536e-26 283-313 BL00766A 21.48 6.063e-24 102-132
615	BL00256	Adipokinetic hormone family proteins.	BL00256 12.28 3.298e-10 746-755
616	BL00319	Amyloidogenic glycoprotein extracellular domain proteins.	BL00319C 17.12 9.053e-09 419-453
617	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030A 14.39 4.429e-09 44-63
618	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030A 14.39 4.429e-09 44-63
620	BL00325	Actin-depolymerizing proteins.	BL00325B 21.66 5.817e-16 77-123
622	BL00972	Ubiquitin carboxyl-terminal hydrolases	BL00972A 11.93 5.500e-19 213-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		family 2 proteins.	231 BL00972D 22.55 2.742e-16 501-526 BL00972B 9.45 1.000e-11 297-307 BL00972C 16.48 3.160e-11 370-385 BL00972E 20.72 7.517e-10 526-548
625	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 6.333e-39 6-45
628	BL00039	DEAD-box subfamily ATP-dependent helicases proteins.	BL00039D 21.67 7.750e-31 478-524 BL00039A 18.44 2.000e-25 198-237 BL00039C 15:63 1.844e-15 327-351 BL00039B 19.19 5.636e-14 242-268
630	PD00306	PROTEIN GLYCOPROTEIN PRECURSOR RE.	PD00306A 10.26 7.000e-12 232-246
631	PD00306	PROTEIN GLYCOPROTEIN PRECURSOR RE.	PD00306A 10.26 7.000e-12 290-304
633	BL00785	5'-nucleotidase proteins.	BL00785C 9.45 3.625e-16 108-122 BL00785E 15.85 4.000e-16 279-295 BL00785A 9.73 6.500e-14 29-40 BL00785B 10.65 5.500e-13 72-86 BL00785D 9.89 4.000e-12 135-145
636	PR00832	PAXILLIN SIGNATURE	PR00832E 14.43 9.901e-14 85-108
637	PR00109	TYROSINE KINASE CATALYTIC DOMAIN SIGNATURE	PR00109B 12.27 6.362e-13 221-240
638	PF00635	MSP (Major sperm protein) domain proteins.	PF00635B 15.84 4.900e-11 463-502
639	PR00860	VERTEBRATE METALLOTHIONEIN SIGNATURE	PR00860B 7.04 1.900e-18 85-99 PR00860C 9.61 1.474e-14 99-109 PR00860A 5.46 1.720e-14 63-76
641	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 4.462e-15 271-284 PD00066 13.92 4.462e-15 299-312 PD00066 13.92 2.800e-14 327-340 PD00066 13.92 2.800e-14 383-396 PD00066 13.92 2.800e-14 411-424 PD00066 13.92 7.000e-14 355-368 PD00066 13.92 8.800e-14 439-452 PD00066 13.92 8.800e-14 495-508 PD00066 13.92 1.500e-13 551-564 PD00066 13.92 7.000e-13 467-480 PD00066 13.92 7.000e-13 523-536 PD00066 13.92 9.500e-13 215-228 PD00066 13.92 9.500e-13 243-256 PD00066 13.92 9.500e-13 579-592 PD00066 13.92 8.615e-10 607-620 PD00066 13.92 1.600e-09 187-200
642	BL00961	Ribosomal protein S28e proteins.	BL00961B 11.24 7.429e-37 67-100 BL00961A 9.90 4.079e-26 42-66
643	BL00585	Ribosomal protein S5 proteins.	BL00585A 28.43 1.391e-40 103-155 BL00585B 18.78 3.250e-30 193-230
647	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 9.400e-10 181-192
648	PR00876	NEMATODE METALLOTHIONEIN SIGNATURE	PR00876C 6.15 9.229e-09 112-126
652	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 5.941e-27 29-68
653	BL00047	Histone H4 proteins.	BL00047A 13.53 1.000e-40 2-41

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			BL00047B 6.51 1.429e-40 41-74 BL00047C 12.18 1.310e-38 74-104
654	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.109e-25 30-69
655	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 3.483e-17 19-63
657	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 8.286e-10 31-40
658	BL00125	Serine/threonine specific protein phosphatases proteins.	BL00125B 21.48 1.000e-40 89-135 BL00125C 19.97 1.000e-40 153-200 BL00125D 33.11 1.000e-40 213-268 BL00125A 14.83 8.941e-38 47-84
659	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-16 492-505 PD00066 13.92 9.308e-15 380-393 PD00066 13.92 6.000e-13 352-365 PD00066 13.92 7.000e-13 240-253 PD00066 13.92 7.500e-13 268-281 PD00066 13.92 7.500e-13 408-421 PD00066 13.92 2.174e-11 464-477 PD00066 13.92 1.000e-10 436-449
660	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.189e-26 29-68
661	BL00795	Involucrin proteins.	BL00795C 17.06 7.882e-15 193-238 BL00795C 17.06 3.797e-13 187-232 BL00795C 17.06 5.014e-13 188-233 BL00795C 17.06 4.506e-12 196-241 BL00795C 17.06 7.896e-12 191-236 BL00795C 17.06 1.667e-11 185-230 BL00795C 17.06 2.000e-11 198-243 BL00795C 17.06 3.778e-11 171-216 BL00795C 17.06 6.111e-11 197-242 BL00795C 17.06 6.444e-11 194-239 BL00795C 17.06 8.000e-11 189-234 BL00795C 17.06 8.556e-11 192-237 BL00795C 17.06 1.733e-10 195-240 BL00795C 17.06 2.779e-10 184-229 BL00795C 17.06 4.035e-10 199-244 BL00795C 17.06 5.081e-10 186-231 BL00795C 17.06 6.965e-10 190-235 BL00795C 17.06 2.700e-09 200-245 BL00795C 17.06 5.800e-09 175-220 BL00795C 17.06 6.500e-09 182-227 BL00795C 17.06 6.600e-09 201-246 BL00795C 17.06 6.600e-09 202-247 BL00795C 17.06 6.600e-09 208-253
662	BL00469	Nucleoside diphosphate kinases proteins.	BL00469 22.22 1.000e-40 149-204
663	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 9.411e-11 331-385
664	BL00601	Tryptophan pentad repeat proteins (IRF family) proteins.	BL00601A 20.29 5.500e-23 7-46 BL00601B 20.92 3.631e-13 69-98
665	BL00082	Extradiol ring-cleavage dioxygenases proteins.	BL00082A 19.07 8.615e-12 49-72
666	DM01537	kw SKI2W SKI2 NUCLEOLAR	DM01537B 21.63 4.073e-37 834-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		HELICASE.	881 DM01537B 21.63 9.750e-21 1669-1716 DM01537A 15.14 8.650e-18 698-718 DM01537A 15.14 6.766e-12 1537-1557
667	DM01537	kw SKI2W SKI2 NUCLEOLAR HELICASE.	DM01537B 21.63 7.923e-38 820-867 DM01537B 21.63 9.750e-21 1655-1702 DM01537A 15.14 8.650e-18 684-704 DM01537A 15.14 6.766e-12 1523-1543
669	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 6.786e-24 849-880 BL00107B 13.31 6.727e-13 916-932
670	BL00299	Ubiquitin domain proteins.	BL00299 28.84 9.735e-27 37-89
671	BL00027	'Homeobox' domain proteins.	BL00027 26.43 6.571e-12 432-475
676	PR00861	ALPHA-LYTIC ENDOPEPTIDASE SERINE PROTEASE (S2A) SIGNATURE	PR00861E 9.88 2.385e-09 206-221
678	BL00225	Crystallins beta and gamma 'Greek key' motif proteins.	BL00225B 18.06 7.517e-24 1805-1840 BL00225B 18.06 8.297e-20 1987-2022 BL00225B 18.06 2.575e-19 1896-1931 BL00225B 18.06 8.200e-19 175-210 BL00225B 18.06 8.200e-19 1698-1733 BL00225B 18.06 4.808e-14 73-108 BL00225B 18.06 4.808e-14 1596-1631 BL00225B 18.06 5.500e-14 2077-2112 BL00225A 13.82 5.829e-12 2043-2064 BL00225A 13.82 3.127e-09 1759-1780
679	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 4.240e-10 169-184 PR00320A 16.74 6.294e-10 169-184
680	BL00243	Integrins beta chain cysteine-rich domain proteins.	BL00243I 31.77 1.143e-11 172-215
681	PR00852	XERODERMA PIGMENTOSUM GROUP D PROTEIN SIGNATURE	PR00852H 5.90 1.000e-29 612-635 PR00852E 8.14 3.769e-27 348-371 PR00852D 11.38 8.875e-27 309-331 PR00852B 11.08 2.800e-25 249-269 PR00852I 17.26 3.500e-25 683-704 PR00852F 11.85 5.909e-24 379-398 PR00852G 16.19 4.462e-23 468-486 PR00852C 8.81 9.143e-23 284-303
682	BL50058	G-protein gamma subunit profile.	BL50058 27.23 1.375e-35 15-63
685	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972A 11.93 7.500e-20 40-58 BL00972D 22.55 3.903e-16 300-325 BL00972B 9.45 1.000e-13 120-130 BL00972E 20.72 5.500e-11 325-347
687	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 4.273e-14 98-138
688	BL00388	Proteasome A-type subunits proteins.	BL00388A 23.14 1.000e-40 8-54 BL00388B 31.38 3.864e-33 66-108 BL00388D 20.71 1.000e-21 153-184 BL00388C 18.79 8.147e-16 126-148
689	PD02796	PROTEIN STEROL CARRIER LIPID-	PD02796B 20.92 1.105e-15 347-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		TRAN.	394
691	PD01572	PHOTOSYSTEM II REACTION CENTRE T PROTEIN PHOTOS.	PD01572 8.77 4.083e-09 1-31
692	BL00028	Zinc finger, C2H2 type, domain proteins.	BL00028 16.07 7.600e-10 488-505
694	BL01013	Oxysterol-binding protein family proteins.	BL01013A 25.14 9.357e-33 527-563 BL01013D 26.81 8.235e-23 814-858 BL01013C 9.97 6.211e-14 615-625 BL01013B 11.33 3.605e-13 592-603
695	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 3.571e-13 164-178 PD00289 9.97 8.650e-11 2147-2161 PD00289 9.97 2.552e-09 23-37
698	PR00161	NICKEL-DEPENDENT HYDROGENASE/B-TYPE CYTOCHROME SIGNATURE	PR00161C 9.51 4.930e-09 282-302
700	PR00749	LYSOZYME G SIGNATURE	PR00749F 13.63 8.636e-13 139-156 PR00749H 8.22 3.681e-12 173-194 PR00749B 16.54 1.419e-11 48-70 PR00749C 7.26 3.060e-11 72-91 PR00749A 10.33 4.815e-10 24-45
703	PR00704	CALPAIN CYSTEINE PROTEASE (C2) FAMILY SIGNATURE	PR00704I 9.52 1.000e-29 476-505 PR00704D 11.05 2.500e-27 132-158 PR00704E 12.55 5.500e-27 162-186 PR00704F 13.61 1.000e-22 187-215 PR00704G 13.87 1.237e-21 317-339 PR00704H 13.38 8.138e-21 367-385 PR00704A 14.68 2.125e-19 27-51 PR00704C 11.88 1.257e-17 96-113 PR00704B 17.94 1.833e-15 72-95
705	PR00859	PROKARYOTE METALLOTHIONEIN SIGNATURE	PR00859C 7.06 2.776e-09 94-111
706	BL00226	Intermediate filaments proteins.	BL00226D 19.10 9.581e-26 369-416 BL00226B 23.86 3.250e-24 203-251 BL00226C 13.23 8.269e-21 268-299 BL00226A 12.77 8.200e-14 103-118
707	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 2.440e-10 2-15
708	BL00361	Ribosomal protein S10 proteins.	BL00361B 18.34 5.101e-10 82-105
709	PR00021	SMALL PROLINE-RICH PROTEIN SIGNATURE	PR00021A 4.31 2.200e-10 2-15
710	BL00514	Fibrinogen beta and gamma chains C-terminal domain proteins.	BL00514C 17.41 8.412e-27 160-197 BL00514E 14.28 8.909e-16 219-236 BL00514H 14.95 1.551e-15 317-342 BL00514G 15.98 7.750e-15 284-314 BL00514D 15.35 4.789e-10 201-214
711	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 8.714e-12 49-90
714	BL00400	LBP / BPI / CETP family proteins.	BL00400C 24.53 6.029e-17 158-202 BL00400D 23.26 2.080e-14 222-259 BL00400A 21.59 1.600e-10 27-59
715	BL01154	RNA polymerases L / 13 to 16 Kd	BL01154B 24.55 5.500e-36 40-76

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		subunits proteins.	BL01154A 18.70 3.000e-22 19-40
716	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 9.786e-32 10-49
717	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 9.206e-14 77-102 BL00215A 15.82 8.412e-10 175-200
719	BL00309	Vertebrate galactoside-binding lectin proteins.	BL00309C 18.65 2.241e-09 62-87
726	BL00687	Aldehyde dehydrogenases glutamic acid proteins.	BL00687E 25.37 7.136e-33 266-316 BL00687D 26.00 5.333e-28 151-198 BL00687B 17.54 3.647e-26 39-81 BL00687C 24.13 6.087e-22 96-133 BL00687F 9.55 2.500e-11 352-363
727	DM01354	kw TRANSCRIPTASE REVERSE II ORF2.	DM01354N 13.17 1.000e-40 129-174 DM01354O 8.73 6.605e-15 180-226
734	PD00301	PROTEIN REPEAT MUSCLE CALCIUM-BI.	PD00301A 10.24 6.400e-09 101-112
735	BL01024	Protein phosphatase 2A regulatory subunit PR55 proteins.	BL01024A 10.26 1.000e-40 22-69 BL01024B 8.91 1.000e-40 86-127 BL01024C 7.80 1.000e-40 146-185 BL01024D 13.22 1.000e-40 185-222 BL01024E 11.96 1.000e-40 222-266 BL01024F 9.42 1.000e-40 266-317 BL01024G 11.09 1.000e-40 317-349 BL01024H 13.88 1.000e-40 389-442
736	PF00913	Trypanosome variant surface glycoprotein.	PF00913D 11.90 7.130e-10 24-51
737	PR00700	PROTEIN TYROSINE PHOSPHATASE SIGNATURE	PR00700D 12.47 2.200e-09 82-101
740	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 1.600e-09 68-83 PR00320A 16.74 7.366e-09 68-83
743	PR00871	DNA NUCLEOTIDYLEXOTRANSFERASE (TDT) SIGNATURE	PR00871G 14.48 8.000e-09 178-201
745	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 2.286e-10 33-42
749	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 5.200e-15 221-246 BL00215A 15.82 7.618e-14 20-45 BL00215A 15.82 8.851e-11 123-148 BL00215B 10.44 9.526e-11 69-82 BL00215B 10.44 7.300e-09 272-285 BL00215B 10.44 8.500e-09 165-178
751	BL50002	Src homology 3 (SH3) domain proteins profile.	BL50002A 14.19 1.000e-14 370-389 BL50002B 15.18 2.200e-10 408-422
752	BL00353	HMG1/2 proteins.	BL00353B 11.47 3.089e-12 390-440
753	PF00622	Domain in SPla and the RYanodine Receptor.	PF00622B 21.00 4.214e-14 47-69
754	BL00211	ABC transporters family proteins.	BL00211A 12.23 8.941e-10 66-78
755	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926F 17.75 7.750e-19 392-415 PR00926C 16.07 5.935e-17 253-274 PR00926D 10.53 2.059e-15 301-320 PR00926E 11.70

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			4.971e-15 344-363 PR00926B 16.07 9.526e-13 210-225 PR00926A 10.41 1.514e-12 197-211
756	BL01187	Calcium-binding EGF-like domain proteins pattern proteins.	BL01187A 9.98 2.125e-12 324-336 BL01187A 9.98 4.789e-11 377-389 BL01187B 12.04 3.057e-10 439-455
757	PF00651	BTB (also known as BR-C/Ttk) domain proteins.	PF00651 15.00 4.429e-10 43-56
758	PR00055	HIV TAT DOMAIN SIGNATURE	PR00055A 8.13 8.855e-09 144-156
759	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 5.304e-11 110-123
760	PR00448	NSF ATTACHMENT PROTEIN SIGNATURE°	PR00448D 12.42 3.455e-27 162-186 PR00448A 10.74 1.273e-22 37-57 PR00448B 16.01 9.379e-21 100-118 PR00448C 11.46 1.000e-20 129-147
765	BL01042	Homoserine dehydrogenase proteins.	BL01042A 13.29 5.909e-11 74-95
766	PR00625	DNAJ PROTEIN FAMILY SIGNATURE	PR00625A 12.84 2.154e-18 26-46 PR00625B 13.48 9.000e-16 57-78
768	BL00762	WHEP-TRS domain proteins.	BL00762A 23.43 8.500e-28 112-149 BL00762B 16.14 3.793e-12 64-78 BL00762A 23.43 6.625e-12 6-43 BL00762C 15.58 4.176e-09 459-472 BL00762D 11.15 9.667e-09 210-220
769	PR00709	AVIDIN SIGNATURE	PR00709A 4.60 1.934e-09 1-20
770	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320C 13.01 1.720e-10 262-277 PR00320A 16.74 2.853e-10 262-277 PR00320C 13.01 4.300e-09 96-111 PR00320B 12.19 5.500e-09 262-277 PR00320A 16.74 6.268e-09 55-70
771	PR00019	LEUCINE-RICH REPEAT SIGNATURE	PR00019B 11.36 8.714e-12 87-101 PR00019A 11.19 1.000e-10 90-104
772	PD02807	APOLIPOPROTEIN E PRECURSOR APO-E GLYCOPROTEIN PLAS.	PD02807C 8.91 6.308e-10 110-159
773	PD02807	APOLIPOPROTEIN E PRECURSOR APO-E GLYCOPROTEIN PLAS.	PD02807C 8.91 6.308e-10 155-204
774	DM00547	1 kw CHROMO BROMODOMAIN SHADOW GLOBAL.	DM00547F 23.43 3.942e-28 943-990 DM00547E 13.94 9.750e-21 652-675 DM00547B 11.28 1.818e-18 518-532 DM00547C 17.30 3.531e-17 546-568 DM00547A 12.38 1.273e-11 497-509 DM00547D 11.60 9.200e-11 622-636
776	PR00779	INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN RECEPTOR SIGNATURE	PR00779F 14.51 5.147e-09 769-792
777	PR00779	INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN RECEPTOR SIGNATURE	PR00779F 14.51 5.147e-09 742-765
778	PR00779	INOSITOL 1,4,5-TRISPHOSPHATE-BINDING PROTEIN RECEPTOR SIGNATURE	PR00779F 14.51 5.147e-09 742-765

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
779	BL01282	BIR repeat proteins.	BL01282B 30.49 2.543e-09 6-45
781	PR00205	CADHERIN SIGNATURE	PR00205B 11.39 3.118e-11 654-672 PR00205B 11.39 8.588e-11 230-248 PR00205B 11.39 8.527e-10 551-569 PR00205B 11.39 4.203e-09 336-354
783	BL00625	Regulator of chromosome condensation (RCC1) proteins.	BL00625B 17.69 2.167e-19 193-227 BL00625A 16.21 5.500e-17 199-228 BL00625B 17.69 1.885e-16 140-174 BL00625B 17.69 2.770e-16 245-279 BL00625A 16.21 9.115e-16 251-280 BL00625A 16.21 6.507e-14 146-175
785	PF00084	Sushi domain proteins (SCR repeat proteins.	PF00084B 9.45 7.188e-10 595-607 PF00084B 9.45 6.400e-09 656-668
786	PF00084	Sushi domain proteins (SCR repeat proteins.	PF00084B 9.45 7.188e-10 595-607 PF00084B 9.45 6.400e-09 656-668
787	BL00826	MARCKS family proteins.	BL00826C 7.63 6.738e-09 203-230
788	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 1.310e-14 36-54 PR00453B 14.65 8.568e-10 75-90
789	PR00102	ORNITHINE CARBAMOYLTRANSFERASE SIGNATURE	PR00102B 14.82 5.418e-09 963-977
790	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030B 7.03 5.500e-11 199-209
791	BL00415	Synapsins proteins.	BL00415N 4.29 9.519e-10 393-437 BL00415N 4.29 2.117e-09 103-147 BL00415N 4.29 3.628e-09 97-141 BL00415N 4.29 5.664e-09 387-431
795	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.091e-36 105-144
799	PF00731	AIR carboxylase.	PF00731C 23.16 7.333e-35 337-380 PF00731B 19.47 7.429e-28 299-336 PF00731A 19.32 6.333e-24 268-297
804	BL00170	Cyclophilin-type peptidyl-prolyl cis-trans isomerase signatur.	BL00170B 20.97 8.071e-09 297-337
805	BL00678	Trp-Asp (WD) repeat proteins proteins.	BL00678 9.67 3.400e-10 378-389 BL00678 9.67 5.800e-10 418-429 BL00678 9.67 8.800e-10 295-306
806	PD01719	PRECURSOR GLYCOPROTEIN SIGNAL RE.	PD01719A 12.89 7.571e-14 290-318
807	PR00320	G-PROTEIN BETA WD-40 REPEAT SIGNATURE	PR00320B 12.19 9.100e-09 451-466
809	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 4.462e-12 564-595
810	PR00453	VON WILLEBRAND FACTOR TYPE A DOMAIN SIGNATURE	PR00453A 12.79 1.310e-14 36-54 PR00453B 14.65 8.568e-10 75-90
814	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.047e-31 16-55
815	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 2.047e-31 16-55
817	PR00193	MYOSIN HEAVY CHAIN SIGNATURE	PR00193D 14.36 5.154e-36 125-154 PR00193E 19.47 3.919e-18 179-208
818	PR00830	ENDOPEPTIDASE LA (LON) SERINE	PR00830A 8.41 9.571e-11 115-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		PROTEASE (S16) SIGNATURE	135
819	BL00126	3'5'-cyclic nucleotide phosphodiesterases proteins.	BL00126C 22.07 7.857e-24 528-569 BL00126E 35.22 3.714e-15 669-724 BL00126D 25.50 1.173e-14 584-623 BL00126B 15.20 1.000e-12 502-514 BL00126A 27.56 3.361e-09 461-498
820	PR00511	TEKTIN SIGNATURE	PR00511B 12.25 8.826e-22 174-195 PR00511A 13.59 7.723e-11 155-172
821	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 2.800e-15 13-36
822	PF00780	Domain found in NIK1-like kinases, mouse citron and yeast ROM.	PF00780I 14.69 4.825e-09 231-261
827	BL00030	Eukaryotic RNA-binding region RNP-1 proteins.	BL00030A 14.39 5.235e-11 144-163
828	BL00326	Tropomyosins proteins.	BL00326D 8.76 9.357e-11 545-586
829	PD02448	TRANSCRIPTION PROTEIN DNA-BINDIN.	PD02448A 9.37 1.000e-40 46-85 PD02448B 10.17 1.000e-40 85-133 PD02448C 13.62 1.000e-40 152-189 PD02448E 11.33 9.000e-30 235-261 PD02448F 14.22 9.654e-25 279-303 PD02448D 11.48 3.659e-18 197-211 PD02448G 10.73 7.857e-16 305-318
830	BL00720	Guanine-nucleotide dissociation stimulators CDC25 family sign.	BL00720B 16.57 4.500e-23 483-507
831	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 6.625e-21 143-174 BL00107B 13.31 4.214e-10 213-229
832	BL00215	Mitochondrial energy transfer proteins.	BL00215A 15.82 5.787e-11 32-57
833	PR00497	NEUTROPHIL CYTOSOL FACTOR P40 SIGNATURE	PR00497A 6.92 4.375e-09 41-59
834	BL00229	Tau and MAP proteins tubulin-binding domain proteins.	BL00229A 23.57 9.565e-10 99-138
835	BL00421	Transmembrane 4 family proteins.	BL00421E 20.97 2.216e-09 1053-1083
836	BL00795	Involucrin proteins.	BL00795B 12.41 7.931e-09 405-445
837	PR00020	MAM DOMAIN SIGNATURE	PR00020A 18.17 1.000e-17 34-53 PR00020B 15.52 5.846e-16 68-85 PR00020D 12.70 2.543e-15 147-162 PR00020C 13.66 3.483e-13 95-107 PR00020E 8.64 6.586e-13 165-179
838	BL50017	Death domain proteins profile.	BL50017B 17.60 6.897e-13 1499-1515
839	PF00850	Histone deacetylase family.	PF00850C 14.55 9.542e-09 1352-1369
840	PF00023	Ank repeat proteins.	PF00023A 16.03 4.500e-12 44-60 PF00023B 14.20 7.923e-11 73-83 PF00023B 14.20 9.000e-10 139-149 PF00023B 14.20 5.500e-09 40-50
842	BL01194	Ribosomal protein L15e proteins.	BL01194B 13.66 1.000e-40 37-85 BL01194C 12.35 9.250e-40 103-138 BL01194A 18.70 7.632e-38

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			2-37 BL01194D 19.02 2.658e-36 139-178
843	BL00610	Sodium:neurotransmitter symporter family proteins.	BL00610A 17.73 1.000e-40 40-90 BL00610B 23.65 1.000e-40 104-154 BL00610C 12.94 1.000e-40 206-258 BL00610E 20.34 1.000e-40 355-398 BL00610F 29.02 1.000e-40 454-509 BL00610D 20.97 6.063e-35 272-325 BL00610G 12.89 8.588e-13 514-537
845	BL00143	Insulinase family, zinc-binding region proteins.	BL00143A 20.91 4.300e-20 94-121 BL00143C 14.16 5.500e-13 245-258 BL00143B 14.41 9.053e-10 141-156
846	PR00543	OESTROGEN RECEPTOR SIGNATURE	PR00543D 10.87 1.355e-09 898-914
847	PR00543	OESTROGEN RECEPTOR SIGNATURE	PR00543D 10.87 1.355e-09 898-914
848	BL00824	Elongation factor 1 beta/beta'/delta chain proteins.	BL00824C 14.58 1.000e-40 129-167 BL00824D 14.04 6.192e-39 167-202 BL00824B 9.21 2.080e-21 96-116 BL00824E 12.49 3.333e-19 210-226 BL00824A 13.78 8.650e-14 19-34
849	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 1.000e-40 12-51
850	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.316e-24 10-49
852	BL01272	Glucokinase regulatory protein family proteins.	BL01272B 19.61 6.870e-30 136-171 BL01272C 11.68 3.314e-25 249-274 BL01272A 6.49 1.231e-18 99-117
853	PD00930	PROTEIN GTPASE DOMAIN ACTIVATION.	PD00930B 33.72 9.341e-20 65-106
854	PD00289	PROTEIN SH3 DOMAIN REPEAT PRESYNA.	PD00289 9.97 6.850e-11 140-154
858	PR00450	RECOVERIN FAMILY SIGNATURE	PR00450C 12.22 3.250e-25 68-90 PR00450B 11.76 8.125e-23 22-42 PR00450D 16.58 8.920e-22 92-112 PR00450E 12.14 1.581e-19 114-133 PR00450G 15.33 5.500e-19 166-187 PR00450F 12.30 4.375e-15 140-156 PR00450A 13.58 1.857e-14 8-23
860	BL00027	'Homeobox' domain proteins.	BL00027 26.43 7.188e-27 74-117
866	BL00477	Alpha-2-macroglobulin family thiolester region proteins.	BL00477L 23.51 7.480e-20 54-87
867	BL01078	Molybdenum cofactor biosynthesis proteins.	BL01078B 14.20 1.621e-20 408-429 BL01078A 10.16 2.000e-13 366-379 BL01078D 5.99 3.455e-11 566-576 BL01078C 10.52 3.793e-11 501-513
868	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 5.800e-24 462-489 BL01177C 17.39 5.333e-19 416-435 BL01177B 13.61 7.840e-16 122-138 BL01177D 17.50 1.900e-15 441-459
869	BL01177	Anaphylatoxin domain proteins.	BL01177E 20.64 5.800e-24 415-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			442 BL01177C 17.39 5.333e-19 369-388 BL01177B 13.61 7.840e-16 122-138 BL01177D 17.50 1.900e-15 394-412
871	BL50007	Phosphatidylinositol-specific phospholipase X-box domain proteins prof.	BL50007A 19.61 1.000e-40 322-368 BL50007D 19.54 1.000e-40 589-631 BL50007B 20.90 6.700e-36 383-421 BL50007E 25.63 9.053e-33 748-785 BL50007C 8.97 5.200e-19 452-469
872	BL00972	Ubiquitin carboxyl-terminal hydrolases family 2 proteins.	BL00972D 22.55 3.250e-17 90-115
874	PR00452	SH3 DOMAIN SIGNATURE	PR00452B 11.65 4.250e-09 370-386
877	BL00741	Guanine-nucleotide dissociation stimulators CDC24 family sign.	BL00741B 14.27 5.500e-13 1343-1366
878	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 2.525e-09 52-85
881	PD02807	APOLIPOPROTEIN E PRECURSOR APO-E GLYCOPROTEIN PLAS.	PD02807E 10.90 4.702e-09 358-407
882	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 7.188e-37 8-47
885	PF00023	Ank repeat proteins.	PF00023A 16.03 8.071e-09 10-26
886	PR00372	BIOPTERIN-DEPENDENT AROMATIC AMINO ACID HYDROXYLASE SIGNATURE	PR00372B 10.30 9.308e-27 225-248 PR00372A 13.39 7.000e-24 134-154 PR00372E 12.62 2.125e-23 360-380 PR00372C 7.90 3.025e-22 289-309 PR00372F 13.09 6.333e-21 395-414 PR00372D 10.22 1.000e-19 329-348
887	BL00301	GTP-binding elongation factors proteins.	BL00301B 20.09 2.800e-24 103-135 BL00301A 12.41 4.316e-13 21-33
888	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 1.667e-09 30-39
889	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 4.906e-26 6-45
890	DM00179	w KINASE ALPHA ADHESION T-CELL.	DM00179 13.97 7.652e-09 113-123
892	BL01022	PTR2 family proton/oligopeptide symporters proteins.	BL01022B 22.19 6.016e-14 72-118 BL01022E 23.51 1.173e-12 472-508 BL01022A 11.58 9.135e-12 42-61 BL01022D 9.42 3.455e-11 199-212
893	PD02407	3-BISPHOSPHOGLYCERATE-INDEPENDENT PHOSPHOGLYCER.	PD02407K 12.59 6.529e-10 360-383
894	PD02407	3-BISPHOSPHOGLYCERATE-INDEPENDENT PHOSPHOGLYCER.	PD02407K 12.59 6.529e-10 360-383
895	PR00237	RHODOPSIN-LIKE GPCR SUPERFAMILY SIGNATURE	PR00237B 13.50 9.100e-14 116-138 PR00237F 13.57 1.360e-13 312-337 PR00237G 19.63 9.069e-13 353-380 PR00237E 13.03 7.120e-12 243-267 PR00237D 8.94 4.150e-11 194-216 PR00237A 11.48 4.375e-11 83-108
896	BL00129	Glycosyl hydrolases family 31 proteins.	BL00129D 16.76 8.258e-26 634-678 BL00129A 26.21 1.720e-25 384-430 BL00129E 22.60 4.857e-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			23 698-734 BL00129C 15.12 1.750e-22 596-624 BL00129B 19.19 5.891e-18 495-522 BL00129F 26.19 7.545e-15 814-852
897	BL00598	Chromo domain proteins.	BL00598 14.45 1.220e-13 9-31
898	BL00518	Zinc finger, C3HC4 type (RING finger), proteins.	BL00518 12.23 6.000e-09 396-405
899	PD01101	INHIBITOR HEAVY CHAIN CHANNEL IN.	PD01101B 21.53 1.000e-40 274-327 PD01101D 24.45 1.000e-40 457-512 PD01101A 18.25 6.268e-23 83-117 PD01101C 12.69 1.237e-16 366-386 PD01101E 6.73 7.750e-12 566-576
900	PR00600	PROTEIN PHOSPHATASE PP2A 55KD REGULATORY SUBUNIT SIGNATURE	PR00600A 11.61 5.979e-09 31-52
901	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 8.116e-31 24-63
903	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 1.509e-11 21-65
906	DM00215	PROLINE-RICH PROTEIN 3.	DM00215 19.43 2.174e-13 539-572 DM00215 19.43 4.750e-12 549-582 DM00215 19.43 9.824e-11 551-584 DM00215 19.43 2.929e-10 548-581 DM00215 19.43 4.054e-10 550-583 DM00215 19.43 5.339e-10 552-585 DM00215 19.43 7.107e-10 544-577
907	PR00988	URIDINE KINASE SIGNATURE	PR00988A 6.39 6.276e-12 314-332
908	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.950e-17 1125-1156
909	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 5.950e-17 1118-1149
910	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 8.560e-13 150-181
911	BL00107	Protein kinases ATP-binding region proteins.	BL00107A 18.39 8.560e-13 150-181
912	PF00856	SET domain proteins.	PF00856A 26.14 4.553e-11 243-280
913	PF00628	PHD-finger.	PF00628 15.84 6.400e-13 197-212
914	PR00962	LETHAL(2) GIANT LARVAE PROTEIN SIGNATURE	PR00962D 10.40 1.000e-27 435-459 PR00962G 15.71 4.086e-26 593-618 PR00962B 11.98 9.122e-26 296-319 PR00962A 13.28 6.143e-22 15-34 PR00962C 8.00 4.000e-21 348-369 PR00962F 12.39 9.769e-21 552-572 PR00962H 13.32 2.636e-20 623-643 PR00962I 11.68 9.786e-20 692-712 PR00962E 8.81 2.915e-18 515-534
915	PR00962	LETHAL(2) GIANT LARVAE PROTEIN SIGNATURE	PR00962D 10.40 1.000e-27 365-389 PR00962G 15.71 4.086e-26 523-548 PR00962A 13.28 6.143e-22 15-34 PR00962C 8.00 4.000e-21 278-299 PR00962F 12.39 9.769e-21 482-502 PR00962H

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			13.32 2.636e-20 553-573 PR00962I 11.68 9.786e-20 622-642 PR00962E 8.81 2.915e-18 445-464
916	BL00134	Serine proteases, trypsin family, histidine proteins.	BL00134A 11.96 5.886e-14 90-107
917	BL00478	LIM domain proteins.	BL00478B 14.79 8.393e-13 211-226 BL00478B 14.79 6.712e-10 271-286
918	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 5.729e-09 973-988
922	BL00150	Acylphosphatase proteins.	BL00150 25.33 1.000e-40 37-84
924	DM00031	IMMUNOGLOBULIN V REGION.	DM00031B 15.41 8.063e-09 79-113
925	BL00072	Acyl-CoA dehydrogenases proteins.	BL00072D 30.08 2.837e-24 280-331 BL00072E 24.12 8.200e-24 368-411 BL00072C 25.30 7.873e-20 226-267 BL00072B 9.48 6.049e-12 183-196
927	BL00237	G-protein coupled receptors proteins.	BL00237C 13.19 1.692e-13 229-256 BL00237A 27.68 6.657e-13 90-130 BL00237D 11.23 9.571e-13 290-307
928	BL01033	Globins profile.	BL01033A 16.94 7.923e-18 25-47 BL01033B 13.81 1.000e-15 93-105
929	BL00216	Sugar transport proteins.	BL00216B 27.64 8.714e-13 203-253
932	BL00415	Synapsins proteins.	BL00415N 4.29 9.519e-10 353-397 BL00415N 4.29 2.117e-09 63-107 BL00415N 4.29 3.628e-09 57-101 BL00415N 4.29 5.664e-09 347-391
933	PD02448	TRANSCRIPTION PROTEIN DNA-BINDIN.	PD02448A 9.37 1.000e-40 46-85 PD02448B 10.17 1.000e-40 85-133 PD02448C 13.62 1.000e-40 152-189 PD02448E 11.33 9.000e-30 223-249 PD02448F 14.22 9.654e-25 267-291 PD02448D 11.48 3.659e-18 197-211 PD02448G 10.73 7.857e-16 293-306
934	DM00191	w SPAC8A4.04C RESISTANCE SPAC8A4.05C DAUNORUBICIN.	DM00191D 13.94 9.083e-10 136-175
935	BL01115	GTP-binding nuclear protein ran proteins.	BL01115A 10.22 4.696e-10 67-111
936	BL00019	Actinin-type actin-binding domain proteins.	BL00019D 15.33 8.138e-14 865-895
937	PR00762	CHLORIDE CHANNEL SIGNATURE	PR00762A 14.22 4.000e-22 183-201 PR00762C 9.29 1.000e-21 268-288 PR00762E 12.07 3.250e-20 520-537 PR00762D 11.29 1.000e-19 470-491 PR00762F 15.12 1.429e-19 538-558 PR00762B 12.12 1.818e-18 214-234 PR00762G 14.13 3.455e-17 577-592
938	BL00027	'Homeobox' domain proteins.	BL00027 26.43 9.500e-25 291-334
939	DM01111	4 kw PHOSPHATASE	DM01111E 17.28 1.568e-10 248-

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
		TRANSFORMING 61K PDF1.	297 DM01111E 17.28 5.168e-10 659-708 DM01111D 16.76 5.263e-09 279-325 DM01111M 10.67 8.674e-09 911-935
940	BL00107	Protein kinases ATP-binding region proteins.	BL00107B 13.31 1.000e-14 293-309 BL00107A 18.39 6.760e-13 229-260
942	BL01160	Kinesin light chain repeat proteins.	BL01160B 19.54 9.832e-11 543-597
943	PD01066	PROTEIN ZINC FINGER ZINC-FINGER METAL-BINDING NU.	PD01066 19.43 3.500e-35 8-47
945	BL00989	Clathrin adaptor complexes small chain proteins.	BL00989B 26.51 1.000e-40 66-117 BL00989A 11.66 1.000e-13 5-19
946	PR00178	FATTY ACID-BINDING PROTEIN SIGNATURE	PR00178D 13.52 9.571e-09 450-469
947	BL00178	Aminoacyl-transfer RNA synthetases class-I proteins.	BL00178B 7.11 4.857e-09 713-724
948	PF00628	PHD-finger.	PF00628 15.84 8.412e-14 201-216
951	BL00216	Sugar transport proteins.	BL00216B 27.64 2.050e-10 180-230
952	PR00926	MITOCHONDRIAL CARRIER PROTEIN SIGNATURE	PR00926F 17.75 4.300e-11 26-49 PR00926F 17.75 6.348e-09 134-157
955	PF00109	Beta-ketoacyl synthase.	PF00109 13.08 2.846e-12 342-357
957	PR00069	ALDO-KETO REDUCTASE SIGNATURE	PR00069A 16.01 8.826e-24 26-51 PR00069B 11.33 1.514e-17 86-105 PR00069C 16.03 8.816e-14 155-173
958	PF00583	Acetyltransferase (GNAT) family.	PF00583A 12.53 5.500e-10 631-642
961	PR00328	GTP-BINDING SAR1 PROTEIN SIGNATURE	PR00328A 10.62 8.740e-10 7-31
962	BL00354	HMG-I and HMG-Y DNA-binding domain proteins (A+T-hook).	BL00354A 3.83 9.438e-10 1489-1499
963	BL00354	HMG-I and HMG-Y DNA-binding domain proteins (A+T-hook).	BL00354A 3.83 9.438e-10 1489-1499
964	BL00027	'Homeobox' domain proteins.	BL00027 26.43 7.188e-27 53-96
965	PF00992	Troponin.	PF00992A 16.67 2.421e-09 581-616
966	PR00515	5-HYDROXYTRYPTAMINE 1F RECEPTOR SIGNATURE	PR00515D 7.91 5.741e-09 13-33
967	BL00579	Ribosomal protein L29 proteins.	BL00579B 21.99 5.065e-21 164-194
970	BL00504	Fumarate reductase / succinate dehydrogenase FAD-binding site proteins.	BL00504C 18.68 2.227e-24 34-59 BL00504D 10.43 7.261e-21 75-93
973	PF00580	UvrD/REP helicase.	PF00580A 13.37 4.720e-09 249-271
974	PR00456	RIBOSOMAL PROTEIN P2 SIGNATURE	PR00456F 5.86 1.000e-10 242-254
975	BL00237	G-protein coupled receptors proteins.	BL00237A 27.68 4.429e-22 99-139
976	BL00031	Nuclear hormones receptors DNA-binding region proteins.	BL00031A 19.55 7.158e-33 60-93 BL00031B 22.25 5.500e-28 94-126
977	PD00066	PROTEIN ZINC-FINGER METAL-BINDI.	PD00066 13.92 8.200e-16 196-209 PD00066 13.92 8.200e-16 336-349 PD00066 13.92 2.385e-15 476-489

SEQ ID NO:	ACCESSION NO.	DESCRIPTION	RESULTS*
			PD00066 13.92 9.308e-15 252-265 PD00066 13.92 2.800e-14 448-461 PD00066 13.92 4.600e-14 392-405 PD00066 13.92 5.200e-14 280-293 PD00066 13.92 4.000e-13 224-237 PD00066 13.92 4.429e-12 308-321 PD00066 13.92 9.571e-12 420-433 PD00066 13.92 6.870e-11 168-181
978	BL00721	Formate--tetrahydrofolate ligase proteins.	BL00721B 13.21 1.000e-40 346-401 BL00721D 13.90 1.000e-40 538-592 BL00721E 13.46 1.000e-40 597-646 BL00721I 18.79 2.500e-40 814-860 BL00721H 21.20 8.239e-39 763-814 BL00721A 15.31 9.719e-32 287-321 BL00721C 16.92 4.000e-30 498-535 BL00721F 15.96 8.232e-27 660-702 BL00721G 7.97 3.017e-10 721-734
981	PD00126	PROTEIN REPEAT DOMAIN TPR NUCLEA.	PD00126A 22.53 2.552e-09 180-201
982	BL00869	Renal dipeptidase proteins.	BL00869C 12.58 3.172e-19 59-95 BL00869E 13.12 9.129e-18 120-157 BL00869J 15.60 6.032e-17 270-310 BL00869H 11.08 1.840e-16 219-242 BL00869G 13.55 2.543e-16 192-214 BL00869F 12.77 7.031e-14 157-192 BL00869I 12.92 3.274e-12 242-270 BL00869D 14.02 5.282e-10 95-124 BL00869B 15.55 9.382e-10 31-61
983	PR00196	ANNEXIN FAMILY SIGNATURE	PR00196F 13.89 2.125e-09 92-108
984	BL00485	Adenosine and AMP deaminase proteins.	BL00485D 30.82 2.427e-10 154-209

* Results include in order: accession number subtype; raw score; p-value; position of signature in amino acid sequence

TABLE 4

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
2	ig	Immunoglobulin domain	3.9e-17	60.3
3	HSP90	Hsp90 protein	0	1548.4
6	tsp_1	Thrombospondin type 1 domain	0.002	22.1
7	7tm_1	7 transmembrane receptor (rhodopsin family)	6.7e-08	27.3
9	PWWP	PWWP domain	8.1e-16	66.0
12	C1q	C1q domain	1.7e-26	101.5
13	C1q	C1q domain	2e-20	81.3
14	Aa_trans	Transmembrane amino acid transporter protein	2.7e-42	153.9
15	E1-E2_ATPase	E1-E2 ATPase	6.3e-124	412.2
16	trypsin	Trypsin	1.2e-87	278.6
17	ig	Immunoglobulin domain	7.6e-12	43.2
18	lectin_c	Lectin C-type domain	0.0003	21.2
20	Alpha_L_fucos	Alpha-L-fucosidase	1.2e-217	736.5

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
22	pkinase	Eukaryotic protein kinase domain	3.3e-87	303.1
23	pkinase	Eukaryotic protein kinase domain	2.7e-85	296.8
24	pkinase	Eukaryotic protein kinase domain	2.7e-85	296.8
25	ank	Ank repeat	5.5e-14	59.9
27	pkinase	Eukaryotic protein kinase domain	1.5e-100	347.4
28	spectrin	Spectrin repeat	4e-57	203.2
29	spectrin	Spectrin repeat	4e-57	203.2
30	WD40	WD domain, G-beta repeat	1.2e-07	38.8
33	rrm	RNA recognition motif.	1.1e-17	72.2
34	rrm	RNA recognition motif.	1.1e-17	72.2
36	7tm_1	7 transmembrane receptor (rhodopsin family)	3e-36	117.3
37	ank	Ank repeat	5.9e-25	96.3
38	SRF-TF	SRF-type transcription factor	1.4e-36	133.9
40	alk_phosphatase	Alkaline phosphatase	0	1034.9
44	zf-C2H2	Zinc finger, C2H2 type	8.6e-103	354.9
45	sugar_tr	Sugar (and other) transporter	3.1e-08	40.3
47	7tm_2	7 transmembrane receptor (Secretin family)	6.4e-79	275.6
50	zf-C2H2	Zinc finger, C2H2 type	1.3e-98	341.0
51	filament	Intermediate filament proteins	1.2e-176	600.3
52	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	2.7e-10	37.7
53	Cadherin_C_ter m	Cadherin cytoplasmic region	1.9e-94	327.2
54	S_100	S-100/ICaBP type calcium binding domain	5.2e-18	73.3
58	inositol_P	Inositol monophosphatase family	5e-13	49.8
59	7tm_1	7 transmembrane receptor (rhodopsin family)	8.8e-46	147.6
60	Kunitz_BPTI	Kunitz/Bovine pancreatic trypsin inhibito	3.7e-47	148.6
62	DAD	DAD family	2.5e-74	260.3
63	MOZ_SAS	MOZ/SAS family	5.9e-133	455.1
64	MOZ_SAS	MOZ/SAS family	1.7e-123	423.6
65	ras	Ras family	9.3e-89	308.3
67	Ham1p_like	Ham1 family	3.7e-49	176.7
68	7tm_1	7 transmembrane receptor (rhodopsin family)	5.2e-39	126.1
70	zf-C2H2	Zinc finger, C2H2 type	1.5e-112	387.3
71	Peptidase_M41	Peptidase family M41	1.2e-110	381.0
72	abhydrolase	alpha/beta hydrolase fold	9.8e-05	26.5
81	K_tetra	K+ channel tetramerisation domain	0.022	-16.8
82	pkinase	Eukaryotic protein kinase domain	5e-49	176.3
84	AAA	ATPases associated with various cellular act	1.3e-77	271.3
85	homeobox	Homeobox domain	1.4e-28	108.3
87	TGF-beta	Transforming growth factor beta like	6.7e-68	210.2
91	mito_carr	Mitochondrial carrier proteins	4.6e-57	198.5
95	adenylatekinase	Adenylate kinase	1.1e-15	60.0
96	ig	Immunoglobulin domain	4.1e-20	69.8
99	CNH	CNH domain	3.4e-120	412.7
100	homeobox	Homeobox domain	7.4e-32	119.3
101	zf-C2H2	Zinc finger, C2H2 type	2.2e-47	170.8
102	zf-C2H2	Zinc finger, C2H2 type	4.4e-89	309.4
103	dynamin	Dynamin family	1.4e-150	513.6
104	lectin_c	Lectin C-type domain	4.2e-15	63.6
105	lectin_c	Lectin C-type domain	4.2e-15	63.6
108	metalthio	Metallothionein	2e-25	97.9

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
112	HSP20	Hsp20/alpha crystallin family	2.6e-20	77.7
115	EF_TS	Elongation factor TS	3.8e-63	221.1
116	sugar_tr	Sugar (and other) transporter	4e-63	223.1
118	catalase	Catalase	0	1158.9
119	UCH	Ubiquitin carboxyl-terminal hydrolase, famil	1e-10	24.4
122	metalthio	Metallothionein	2.8e-25	97.4
125	adh_short	short chain dehydrogenase	1.6e-45	164.6
126	KRAB	KRAB box	7.9e-25	95.9
127	G-alpha	G-protein alpha subunit	1e-249	843.0
128	mito_carr	Mitochondrial carrier proteins	2e-65	227.2
131	EF1BD	EF-1 guanine nucleotide exchange domain	4.9e-53	189.6
132	GYF	GYF domain	4.9e-28	106.6
133	GYF	GYF domain	4.9e-28	106.6
134	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	2.1e-33	119.1
135	pkinase	Eukaryotic protein kinase domain	3.3e-86	299.8
136	ank	Ank repeat	2.2e-29	111.1
137	IL8	Small cytokines (intecrine/chemokine), inter	3.1e-18	65.2
139	pyridoxal_deC	Pyridoxal-dependent decarboxylase conse	0.00011	19.0
140	cadherin	Cadherin domain	1.3e-88	307.8
142	efhand	EF hand	5.7e-33	123.0
143	Acyltransferase	Acyltransferase	2e-29	111.2
146	cytochrome_c	Cytochrome c	1.7e-33	124.7
147	pkinase	Eukaryotic protein kinase domain	2.3e-86	300.3
148	PDZ	PDZ domain (Also known as DHR or GLGF).	1.7e-09	45.0
149	aldo_ket_red	Aldo/keto reductase family	7.4e-189	640.8
150	homeobox	Homeobox domain	3.2e-08	38.7
151	PseudoU_synth_1	tRNA pseudouridine synthase	4.7e-57	203.0
152	abhydrolase	alpha/beta hydrolase fold	1.7e-31	118.0
153	PDZ	PDZ domain (Also known as DHR or GLGF).	1.1e-09	45.6
156	PHD	PHD-finger	7.6e-15	62.8
157	fn3	Fibronectin type III domain	0.015	21.9
158	homeobox	Homeobox domain	2.7e-27	104.1
160	PWI	PWI domain	3.9e-24	93.6
162	DnaJ	DnaJ domain	2e-06	34.8
164	Cbl_N	CBL proto-oncogene N-terminal domain	8e-117	401.5
166	metalthio	Metallothionein	3.1e-26	100.6
167	LRR	Leucine Rich Repeat	0.00069	26.3
169	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	5.3e-180	611.4
170	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	5.3e-180	611.4
171	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	1e-149	510.8
173	homeobox	Homeobox domain	1.5e-29	111.6
174	FYVE	FYVE zinc finger	7.4e-28	103.8
175	GRIP	GRIP domain	3.9e-08	40.5
182	pkinase	Eukaryotic protein kinase domain	3.4e-71	250.0
185	CAP_GLY	CAP-Gly domain	5.6e-51	182.8
186	TBC	TBC domain	2.2e-50	180.8
187	TBC	TBC domain	2.2e-50	180.8

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
188	PDZ	PDZ domain (Also known as DHR or GLGF).	4e-13	57.0
189	Kelch	Kelch motif	5.2e-106	365.6
190	Tropomyosin	Tropomyosins	3.8e-171	535.4
192	Rieske	Rieske [2Fe-2S] domain	0.0016	18.5
199	ig	Immunoglobulin domain	5.9e-19	66.1
202	EGF	EGF-like domain	3.4e-54	193.5
203	trefoil	Trefoil (P-type) domain	1e-24	95.5
204	TBC	TBC domain	8.5e-38	139.0
205	efhand	EF hand	0.0096	22.6
206	ISK_Channel	Slow voltage-gated potassium channel	0.0031	8.1
207	trefoil	Trefoil (P-type) domain	2.9e-48	173.7
209	Ribosomal_S13	Ribosomal protein S13/S18	1.2e-78	274.7
210	hemopexin	Hemopexin	1.3e-62	221.5
213	TBC	TBC domain	2.5e-48	174.0
215	Basic	Myogenic Basic domain	4.3e-50	179.8
216	Ribosomal_L24	KOW motif	8.2e-23	89.2
222	fn3	Fibronectin type III domain	7.3e-141	481.4
223	cofilin_ADF	Cofilin/tropomyosin-type actin-binding pr	9.3e-47	168.8
224	efhand	EF hand	6.1e-06	33.2
225	Pterin_4a	Pterin 4 alpha carbinolamine dehydratase	9.3e-42	152.1
228	ABC_tran	ABC transporter	4.1e-110	379.2
234	E1_DerP2_DerF 2	E1 family	3.7e-90	312.9
235	E1_DerP2_DerF 2	E1 family	1.6e-48	174.6
237	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	1.7e-25	98.1
238	Opiods_neurope p	Vertebrate endogenous opioids neurope	1.8e-159	543.2
239	eIF-5a	Eukaryotic initiation factor 5A hypusine	5.9e-104	358.8
240	Amino_oxidase	Flavin containing amine oxidase	2.5e-11	37.8
243	zf-C2H2	Zinc finger, C2H2 type	2.1e-99	343.6
244	Band_7	SPFH domain / Band 7 family	2.3e-53	190.7
245	ank	Ank repeat	1.6e-88	307.5
246	zf-C2H2	Zinc finger, C2H2 type	6.7e-49	175.9
247	actin	Actin	2.3e-42	140.3
248	ER_lumen_recep t	ER lumen protein retaining receptor	2.4e-155	529.5
250	PMP22_Claudin	PMP-22/EMP/MP20/Claudin family	2.2e-38	140.9
252	Collagen	Collagen triple helix repeat (20 copies)	1.4e-13	58.6
255	C2	C2 domain	0.052	7.8
257	CAP_GLY	CAP-Gly domain	1.4e-20	81.8
260	WD40	WD domain, G-beta repeat	9.9e-62	218.5
261	WD40	WD domain, G-beta repeat	9.9e-62	218.5
262	WD40	WD domain, G-beta repeat	9.9e-62	218.5
263	cofilin_ADF	Cofilin/tropomyosin-type actin-binding pr	7.8e-21	82.6
264	Ribosomal_L14	Ribosomal protein L14p/L23e	9.2e-10	40.6
265	SAPA	Saposin A-type domain	4.4e-27	103.4
266	SAPA	Saposin A-type domain	4.4e-27	103.4
267	ABC_tran	ABC transporter	9.5e-39	142.2
269	Ribosomal_L14	Ribosomal protein L14p/L23e	6.2e-62	219.2
270	abhydrolase	alpha/beta hydrolase fold	0.042	-3.3
272	ras	Ras family	4.3e-87	302.8

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
273	rrm	RNA recognition motif.	0.074	14.6
275	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	2.5e-41	146.4
276	ras	Ras family	1.1e-67	238.3
277	UCH	Ubiquitin carboxyl-terminal hydrolase, famil	1.2e-147	503.9
278	START	START domain	3.2e-09	44.1
279	WD40	WD domain, G-beta repeat	1.8e-27	104.7
282	G-patch	G-patch domain	7.8e-22	86.0
287	Anti_proliferat	BTG1 family	1.2e-101	351.0
289	KRAB	KRAB box	7.1e-21	82.8
293	7tm_3	7 transmembrane receptor	3.3e-73	256.6
295	SET	SET domain	5e-30	113.2
296	Pyridox_oxidase	Pyridoxamine 5'-phosphate oxidase	1.3e-76	268.0
297	rrm	RNA recognition motif.	5.4e-45	162.9
298	Ubie_methyltran	ubiE/COQ5 methyltransferase family	6.3e-05	-96.3
299	Ubie_methyltran	ubiE/COQ5 methyltransferase family	0.0024	-118.1
301	Cyt_reductase	FAD/NAD-binding Cytochrome reductase	7.7e-61	215.5
302	G-patch	G-patch domain	3.1e-14	60.7
307	7tm_1	7 transmembrane receptor (rhodopsin family)	7.7e-43	138.2
308	PH	PH domain	0.0015	17.8
310	7tm_1	7 transmembrane receptor (rhodopsin family)	1.4e-84	270.8
311	Rhodanese	Rhodanese-like domain	3.3e-64	226.7
312	tubulin	Tubulin/FtsZ family	4.9e-286	963.6
314	SURF4	SURF4 family	1.2e-199	676.6
325	IMS	impB/mucB/samB family	2e-58	207.5
327	cadherin	Cadherin domain	4.3e-91	316.0
329	NAC	NAC domain	2.1e-28	107.8
330	IP_trans	Phosphatidylinositol transfer protein	6.5e-98	338.7
332	TFIIS	Transcription factor S-II (TFIIS)	8.8e-05	29.3
337	zf-C2H2	Zinc finger, C2H2 type	3.6e-61	216.6
340	AIRS	AIR synthase related protein	4e-32	120.2
343	annexin	Annexin	4.6e-80	279.4
346	Stathmin	Stathmin family	1.8e-90	314.0
347	Ribosomal_L16	Ribosomal protein L16	4.6e-09	34.9
348	lactamase_B	Metallo-beta-lactamase superfamily	0.012	-6.0
351	efhand	EF hand	2.5e-14	61.0
353	lectin_c	Lectin C-type domain	1.3e-05	32.1
354	WD40	WD domain, G-beta repeat	2.2e-18	74.5
360	lipocalin	Lipocalin / cytosolic fatty-acid binding pr	6.3e-10	38.3
362	Acetyltransf	Acetyltransferase (GNAT) family	0.0019	24.9
365	tRNA-synt_1	tRNA synthetases class I (I, L, M and V)	4.6e-185	628.2
366	Sulfatase	Sulfatase	6.1e-228	770.6
368	START	START domain	3.8e-11	50.5
369	pkinase	Eukaryotic protein kinase domain	2.4e-10	41.3
370	ACBP	Acyl CoA binding protein	4.4e-56	199.7
371	pkinase	Eukaryotic protein kinase domain	1.6e-94	327.5
373	EGF	EGF-like domain	2.6e-12	54.3
375	zf-C2H2	Zinc finger, C2H2 type	8.2e-64	225.4
377	KRAB	KRAB box	3.7e-27	103.7
379	SET	SET domain	7.3e-61	215.6
380	Glyco_transf_8	Glycosyl transferase family 8	0.0028	-40.1
381	zf-C2H2	Zinc finger, C2H2 type	4.3e-06	33.7
383	Glyco_transf_8	Glycosyl transferase family 8	0.0028	-40.1

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
384	RasGEF	RasGEF domain	8.1e-43	155.7
385	TBC	TBC domain	0.017	-66.6
389	Glycos_transf_2	Glycosyl transferases	1.3e-15	65.3
390	Na_Ca_Ex	Sodium/calcium exchanger protein	3.9e-105	362.7
391	fn3	Fibronectin type III domain	4.1e-102	352.6
392	fn3	Fibronectin type III domain	3.4e-45	163.6
393	fn3	Fibronectin type III domain	3.4e-45	163.6
394	ldl_recept_b	Low-density lipoprotein receptor repeat	7.1e-49	175.8
395	Ribosomal_L30	Ribosomal protein L30p/L7e	0.0023	16.0
396	Oxysterol_BP	Oxysterol-binding protein	1.5e-94	327.5
397	RDS_ROM1	Peripherin/rom-1	2.9e-33	123.9
399	lactamase_B	Metallo-beta-lactamase superfamily	3.4e-39	143.6
402	F-box	F-box domain.	0.0002	28.1
403	CLP_protease	Clp protease	4.8e-64	226.2
405	Ribosomal_L35Ae	Ribosomal protein L35Ae	6e-77	269.0
406	LIM	LIM domain containing proteins	0.00021	20.7
410	tRNA-synt_1c	tRNA synthetases class I (E and Q)	1e-236	799.8
411	NTP_transf_2	Nucleotidyltransferase domain	3.9e-16	67.0
412	DEAD	DEAD/DEAH box helicase	0.00016	17.2
414	DUF94	Domain of unknown function DUF94	0.00011	26.9
415	tubulin	Tubulin/FtsZ family	4.5e-289	973.7
420	SET	SET domain	3.3e-57	203.5
421	WD40	WD domain, G-beta repeat	6.1e-29	109.6
423	zf-C2H2	Zinc finger, C2H2 type	1.5e-39	144.9
424	pkinase	Eukaryotic protein kinase domain	8.9e-75	261.8
428	LIM	LIM domain containing proteins	1.8e-34	126.7
431	kazal	Kazal-type serine protease inhibitor domain	3.7e-18	73.8
432	SH2	Src homology domain 2	1.4e-67	198.4
433	zf-C2H2	Zinc finger, C2H2 type	2.8e-144	492.7
434	ras	Ras family	0.012	-106.8
436	E1-E2_ATPase	E1-E2 ATPase	1.6e-117	391.0
437	RNA_pol_A	RNA polymerase alpha subunit	0	1077.7
438	PHD	PHD-finger	1.6e-11	51.7
439	lectin_c	Lectin C-type domain	4.7e-30	113.3
440	zf-C2H2	Zinc finger, C2H2 type	1.1e-65	231.6
441	arrestin	Arrestin (or S-antigen)	2.9e-254	858.1
442	aminotran_3	Aminotransferases class-III pyridoxal-pho	8.2e-80	231.1
443	UCH-1	Ubiquitin carboxyl-terminal hydrolases famil	8.5e-12	52.6
444	CTF_NFI	CTF/NF-I family	2.6e-277	934.6
451	T-box	T-box	3.8e-117	402.6
453	Rieske	Rieske [2Fe-2S] domain	2.6e-13	57.7
454	zf-C2H2	Zinc finger, C2H2 type	3.9e-64	226.5
456	homeobox	Homeobox domain	2.8e-08	38.9
459	ig	Immunoglobulin domain	2.6e-20	70.5
460	Hydrolase	haloacid dehalogenase-like hydrolase	4e-25	96.9
462	rve	Integrase core domain	1.6e-13	50.7
466	CH	Calponin homology (CH) domain	2.4e-17	71.1
467	CH	Calponin homology (CH) domain	2.4e-17	71.1
468	Sterol_desat	Sterol desaturase	7.5e-38	139.2
469	pro_isomerase	Cyclophilin type peptidyl-prolyl cis-tr	2.6e-63	220.9
470	Peptidase_M24	metallopeptidase family M24	6e-08	28.1
471	PDZ	PDZ domain (Also known as DHR or GLGF).	5.4e-129	441.9

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
472	myb_DNA-binding	Myb-like DNA-binding domain	3.6e-06	33.9
473	ZZ	Zinc finger present in dystrophin, CB	0.012	20.0
474	EF1G_domain	Elongation factor 1 gamma, conserved doma	6.3e-88	305.5
475	Ribosomal_L31e	Ribosomal protein L31e	6.1e-66	232.5
476	C1q	C1q domain	2.5e-75	263.7
477	SH3	SH3 domain	1.1e-12	55.6
478	MoaA_NifB_PqqE	moaA / nifB / pqqE family	0.002	-17.7
479	FYVE	FYVE zinc finger	9.3e-21	78.6
480	DNA_pol_A	DNA polymerase family A	2.3e-46	167.4
482	adh_short	short chain dehydrogenase	1.2e-62	221.6
483	ank	Ank repeat	1.3e-17	71.9
484	IMS	impB/mucB/samB family	2.2e-83	290.5
486	TIR	TIR domain	3.2e-19	67.8
487	FMO-like	Flavin-binding monooxygenase-like	0	1425.5
488	I_LWEQ	I/LWEQ domain	9.5e-101	341.0
495	homeobox	Homeobox domain	3.6e-06	30.8
497	pkinase	Eukaryotic protein kinase domain	2.3e-166	566.1
499	fn3	Fibronectin type III domain	2.5e-237	801.8
501	LRR	Leucine Rich Repeat	9.3e-31	115.6
502	RGS	Regulator of G protein signaling domain	0.041	11.9
503	filament	Intermediate filament proteins	1e-142	487.5
505	fn3	Fibronectin type III domain	1.3e-100	347.7
506	HECT	HECT-domain (ubiquitin-transferase).	1e-13	59.0
507	Ribosomal_L7Ae	Ribosomal protein L7Ae	5.7e-26	99.7
508	WD40	WD domain, G-beta repeat	0.063	19.8
509	WD40	WD domain, G-beta repeat	0.063	19.8
510	WD40	WD domain, G-beta repeat	2.1e-42	154.3
511	pkinase	Eukaryotic protein kinase domain	2.3e-86	300.4
512	G-gamma	GGL domain	1.9e-08	34.3
513	SH3	SH3 domain	3e-06	34.2
515	HTH_AraC	Bacterial regulatory helix-turn-helix protei	3.9e-27	103.6
516	zf-C2H2	Zinc finger, C2H2 type	1.7e-34	128.0
517	S1	S1 RNA binding domain	6.1e-58	205.9
518	pkinase	Eukaryotic protein kinase domain	1.8e-75	264.2
525	cadherin	Cadherin domain	2e-80	280.6
528	zf-C2H2	Zinc finger, C2H2 type	4e-70	246.4
529	neur_chan	Neurotransmitter-gated ion-channel	5.8e-222	750.8
531	RhoGEF	RhoGEF domain	3.5e-44	160.2
532	myosin_head	Myosin head (motor domain)	0	1494.5
533	LRR	Leucine Rich Repeat	8.3e-15	62.6
535	Sec7	Sec7 domain	5.1e-92	319.1
536	homeobox	Homeobox domain	4.8e-05	26.4
539	actin	Actin	2.4e-100	330.6
542	ank	Ank repeat	1.9e-35	131.2
544	zf-CCCH	Zinc finger C-x8-C-x5-C-x3-H type	2.8e-10	41.7
546	DSPc	Dual specificity phosphatase, catalytic doma	2.4e-40	147.4
547	HMG_CoA_synt	Hydroxymethylglutaryl-coenzyme A synthas	0	1250.8
549	laminin_G	Laminin G domain	3.3e-76	266.6
551	PHD	PHD-finger	0.008	9.3
552	PDZ	PDZ domain (Also known as DHR or	0.0017	25.0

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
		GLGF).		
555	WW	WW domain	1.3e-24	95.3
558	kinesin	Kinesin motor domain	1.8e-176	599.7
559	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	0.00085	16.5
563	efhand	EF hand	7.9e-11	49.4
567	PH	PH domain	7.8e-06	25.9
568	PH	PH domain	3.1e-39	143.8
569	Hist_deacetyl	Histone deacetylase family	5.2e-106	365.6
570	PDZ	PDZ domain (Also known as DHR or GLGF).	3.4e-20	80.5
571	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1e-16	58.5
573	ubiquitin	Ubiquitin family	1.4e-08	31.1
574	FH2	Formin Homology 2 Domain	1.3e-110	380.9
576	serpin	Serpins (serine protease inhibitors)	4.3e-146	496.4
579	zf-C2H2	Zinc finger, C2H2 type	5.7e-76	265.8
580	pkinase	Eukaryotic protein kinase domain	6.9e-79	275.5
581	RhoGAP	RhoGAP domain	4.4e-53	189.8
582	Ribosomal_L7Ae	Ribosomal protein L7Ae	0.028	1.0
584	kazal	Kazal-type serine protease inhibitor domain	2.2e-52	187.4
585	LRR	Leucine Rich Repeat	4.4e-28	106.7
586	PHD	PHD-finger	3.8e-12	53.8
588	GTP1_OBG	GTP1/OBG family	1.1e-62	215.2
590	Collagen	Collagen triple helix repeat (20 copies)	8e-42	152.4
591	lys	C-type lysozyme/alpha-lactalbumin family	1.6e-31	116.4
596	ACBP	Acyl CoA binding protein	0.0022	-9.4
597	SNF2_N	SNF2 and others N-terminal domain	3.7e-98	339.5
600	KRAB	KRAB box	1.3e-29	111.8
606	LRR	Leucine Rich Repeat	1e-05	32.5
607	LRR	Leucine Rich Repeat	1e-05	32.5
608	WD40	WD domain, G-beta repeat	5.3e-23	89.8
610	cpn60_TCP1	TCP-1/cpn60 chaperonin family	1.7e-237	802.4
613	THF_DHG_CYH	Tetrahydrofolate dehydrogenase/cyclohydro	4.9e-173	588.3
617	rrm	RNA recognition motif.	4e-14	60.4
618	rrm	RNA recognition motif.	4e-14	60.4
620	cofilin_ADF	Cofilin/tropomyosin-type actin-binding pr	3e-06	34.2
621	Nop	Putative snoRNA binding domain	6.1e-95	328.8
622	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	5.8e-21	83.1
625	zf-C2H2	Zinc finger, C2H2 type	2.5e-124	426.4
628	DEAD	DEAD/DEAH box helicase	2.5e-68	219.0
632	GST	Glutathione S-transferases.	4.8e-26	89.0
633	5_nucleotidase	5'-nucleotidase	6.6e-248	837.0
636	LIM	LIM domain containing proteins	1.6e-88	307.5
637	pkinase	Eukaryotic protein kinase domain	1.5e-73	257.8
638	MSP_domain	MSP (Major sperm protein) domain	8.4e-09	42.7
639	metalthio	Metallothionein	2e-24	94.6
641	zf-C2H2	Zinc finger, C2H2 type	6.1e-114	391.9
642	Ribosomal_S28e	Ribosomal protein S28e	9.3e-48	172.1
643	Ribosomal_S5	Ribosomal protein S5	8.3e-87	301.8
646	PHD	PHD-finger	0.00025	23.1
647	WD40	WD domain, G-beta repeat	1.5e-22	88.4

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
648	Lipase_GDSL	Lipase/Acylhydrolase with GDSL-like motif	0.015	2.2
652	zf-C2H2	Zinc finger, C2H2 type	4.1e-146	498.8
653	histone	Core histone H2A/H2B/H3/H4	1.2e-10	48.8
654	zf-C2H2	Zinc finger, C2H2 type	1.9e-87	303.9
655	ras	Ras family	6.4e-77	269.0
657	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	5.3e-13	46.4
658	STphosphatase	Ser/Thr protein phosphatase	2.6e-182	619.1
659	zf-C2H2	Zinc finger, C2H2 type	1.3e-92	321.1
660	zf-C2H2	Zinc finger, C2H2 type	1.5e-85	297.6
662	NDK	Nucleoside diphosphate kinases	1.4e-119	410.7
664	IRF	Interferon regulatory factor transcription f	7e-20	79.5
665	4HPPD_C	4-hydroxyphenylpyruvate dioxygenase C term	1.4e-16	68.5
666	DEAD	DEAD/DEAH box helicase	4.8e-74	237.1
667	DEAD	DEAD/DEAH box helicase	2.9e-70	225.1
669	pkinase	Eukaryotic protein kinase domain	6.1e-93	322.2
671	homeobox	Homeobox domain	0.018	16.5
678	crystall	Beta/Gamma crystallin	4.7e-106	365.8
679	WD40	WD domain, G-beta repeat	1.9e-06	34.9
680	Keratin_B2	Keratin, high sulfur B2 protein	4.1e-06	15.9
682	G-gamma	GGL domain	8.5e-33	117.9
685	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	1.4e-29	111.7
686	Acetyltransf	Acetyltransferase (GNAT) family	6.6e-10	46.4
687	7tm_1	7 transmembrane receptor (rhodopsin family)	4.6e-15	50.0
688	proteasome	Proteasome A-type and B-type	6.5e-64	225.7
689	SCP2	SCP-2 sterol transfer family	6.2e-37	136.1
690	TS-N	TS-N domain	0.041	20.1
692	zf-C2H2	Zinc finger, C2H2 type	9.9e-60	211.9
693	zf-MYND	MYND finger	0.038	5.5
694	Oxysterol_BP	Oxysterol-binding protein	3.9e-133	455.7
695	PDZ	PDZ domain (Also known as DHR or GLGF).	1.3e-30	115.1
703	Peptidase_C2	Calpain family cysteine protease	2.3e-175	596.0
706	filament	Intermediate filament proteins	7.2e-107	368.5
710	fibrinogen_C	Fibrinogen beta and gamma chains, C-term	7e-80	278.0
711	SH2	Src homology domain 2	2.3e-65	192.1
712	ATP-synt_DE	ATP synthase, Delta/Epsilon chain	0.00062	19.0
713	ARID	ARID DNA binding domain	2e-17	71.3
714	LBP_BPI_CETP	LBP / BPI / CETP family	8.6e-34	125.7
715	RNA_pol_L	RNA polymerases L / 13 to 16 kDa subunit	4.8e-49	176.3
716	KRAB	KRAB box	1.3e-42	155.0
717	mito_carr	Mitochondrial carrier proteins	4.8e-38	133.3
719	Gal-bind_lectin	Vertebrate galactoside-binding lectin	1.5e-25	90.2
726	aldedh	Aldehyde dehydrogenase family	1.3e-119	410.8
728	Glycos_transf_2	Glycosyl transferases	4e-21	83.6
734	ELM2	ELM2 domain	2e-34	127.8
735	PR55	Protein phosphatase 2A regulatory subunit PR	0	1038.2
737	DSPc	Dual specificity phosphatase, catalytic doma	4e-14	60.4
740	WD40	WD domain, G-beta repeat	5.6e-14	59.9
745	zf-C3HC4	Zinc finger, C3HC4 type (RING	3.8e-13	46.9

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
		finger)		
749	mito_carr	Mitochondrial carrier proteins	4.5e-67	232.8
750	DUF27	Domain of unknown function DUF27	4.5e-12	53.5
751	SH3	SH3 domain	3.6e-17	70.5
752	HMG_box	HMG (high mobility group) box	8.6e-13	55.9
753	SPRY	SPRY domain	5.9e-05	23.3
754	GTP_CDC	Cell division protein	7.5e-153	521.2
755	mito_carr	Mitochondrial carrier proteins	3e-88	305.4
756	TSPN	Thrombospondin N-terminal -like domains	8.1e-58	205.5
757	BTB	BTB/POZ domain	5.7e-23	89.7
759	zf-C2H2	Zinc finger, C2H2 type	1.2e-12	55.4
760	NSF	NSF attachment protein	6.4e-127	435.1
762	Ribosomal_S14	Ribosomal protein S14p/S29e	2.1e-06	24.8
765	ThiF_family	ThiF family	1.7e-39	144.6
766	DnaJ	DnaJ domain	3.9e-36	133.5
768	tRNA-synt_2b	tRNA synthetase class II	9.1e-81	281.7
769	ldl_recept_a	Low-density lipoprotein receptor domain	0	1404.5
770	WD40	WD domain, G-beta repeat	2e-21	84.6
771	LRR	Leucine Rich Repeat	3.8e-06	33.9
774	SNF2_N	SNF2 and others N-terminal domain	5.5e-99	342.3
776	VPS9	Vacuolar sorting protein 9 (VPS9) domain	1.1e-30	115.4
777	VPS9	Vacuolar sorting protein 9 (VPS9) domain	1.1e-30	115.4
778	VPS9	Vacuolar sorting protein 9 (VPS9) domain	1.1e-30	115.4
779	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	3.1e-08	31.0
781	cadherin	Cadherin domain	5.6e-113	388.7
783	HECT	HECT-domain (ubiquitin-transferase).	4.2e-31	116.8
785	sushi	Sushi domain (SCR repeat)	1.8e-60	214.3
786	sushi	Sushi domain (SCR repeat)	1.8e-60	214.3
788	vwa	von Willebrand factor type A domain	1.9e-52	187.7
790	rrm	RNA recognition motif.	2.8e-20	80.8
791	Collagen	Collagen triple helix repeat (20 copies)	0.00097	9.7
792	pkinase	Eukaryotic protein kinase domain	0.023	12.4
795	zf-C2H2	Zinc finger, C2H2 type	6.5e-95	328.7
796	adh_short	short chain dehydrogenase	4.1e-05	-7.3
799	SAICAR_synt	SAICAR synthetase	6e-125	428.5
805	WD40	WD domain, G-beta repeat	4e-65	229.8
806	ZU5	ZU5 domain	4.7e-37	136.5
807	WD40	WD domain, G-beta repeat	0.016	21.8
808	WD40	WD domain, G-beta repeat	0.0041	23.8
809	pkinase	Eukaryotic protein kinase domain	2e-31	117.2
810	vwa	von Willebrand factor type A domain	1.9e-52	187.7
814	zf-C2H2	Zinc finger, C2H2 type	4.5e-83	289.4
815	zf-C2H2	Zinc finger, C2H2 type	6e-74	259.1
817	myosin_head	Myosin head (motor domain)	1.5e-176	599.9
818	GSPII_E	Bacterial type II secretion system protein	0.012	11.5
819	PDEase	3'5'-cyclic nucleotide phosphodiesterase	1.1e-74	215.5
821	PH	PH domain	0.00025	20.5
822	CNH	CNH domain	0.00015	-24.7
827	rrm	RNA recognition motif.	1.5e-06	35.2

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
829	HMG_box	HMG (high mobility group) box	7.8e-34	125.8
830	RasGEF	RasGEF domain	2.2e-102	353.5
831	CNH	CNH domain	3e-118	406.2
832	mito_carr	Mitochondrial carrier proteins	3.7e-37	130.3
833	PX	PX domain	2.7e-19	77.5
837	Y_phosphatase	Protein-tyrosine phosphatase	1.6e-263	888.8
838	ank	Ank repeat	2.4e-270	911.5
840	ank	Ank repeat	5.8e-38	139.6
842	Ribosomal_L15e	Ribosomal L15	4.8e-131	448.8
843	SNF	Sodium:neurotransmitter symporter family	0	1201.8
845	Peptidase_M16	Insulinase (Peptidase family M16)	4.7e-67	236.2
848	EF1BD	EF-1 guanine nucleotide exchange domain	2.2e-56	200.7
849	zf-C2H2	Zinc finger, C2H2 type	1.5e-122	420.5
850	zf-C2H2	Zinc finger, C2H2 type	2e-67	237.4
852	SIS	SIS domain	3.8e-30	113.6
853	RhoGAP	RhoGAP domain	1.1e-37	138.6
854	PDZ	PDZ domain (Also known as DHR or GLGF).	5.1e-10	46.7
856	ACOX	Acyl-CoA oxidase	9.1e-263	886.3
858	efhand	EF hand	2.4e-18	74.4
860	homeobox	Homeobox domain	4e-22	86.9
862	TFIIF_beta	Transcription initiation factor IIF, beta	2.2e-134	459.8
866	A2M	Alpha-2-macroglobulin family	4.9e-21	70.9
867	MoCF_biosynth	Molybdenum cofactor biosynthesis protei	5.8e-205	694.3
868	EGF	EGF-like domain	4.1e-22	86.9
869	EGF	EGF-like domain	1.1e-22	88.8
871	PI-PLC-X	Phosphatidylinositol-specific phospholipase	7.2e-95	328.6
872	UCH-2	Ubiquitin carboxyl-terminal hydrolase family	1.1e-20	82.1
874	SH3	SH3 domain	2.2e-14	61.2
877	SH3	SH3 domain	8.6e-90	311.7
882	KRAB	KRAB box	6.9e-45	162.6
885	ank	Ank repeat	7.1e-07	36.3
886	biopterin_H	Biopterin-dependent aromatic amino acid h	0	988.3
887	GTP_EFTU	Elongation factor Tu family	4.9e-129	437.5
888	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	1.6e-14	51.4
889	zf-C2H2	Zinc finger, C2H2 type	3.7e-92	319.6
890	ig	Immunoglobulin domain	3.8e-06	24.8
892	PTR2	POT family	9.5e-48	163.0
893	Sulfatase	Sulfatase	3.5e-78	273.2
894	Sulfatase	Sulfatase	3.5e-78	273.2
895	7tm_1	7 transmembrane receptor (rhodopsin family)	4.5e-51	164.4
896	Glyco_hydro_31	Glycosyl hydrolases family 31	0	1277.3
897	chromo	'chromo' (CHRromatin Organization MOdifier)	3.9e-06	26.0
898	Cbl_N	CBL proto-oncogene N-terminal domain	1.2e-273	922.4
899	vwa	von Willebrand factor type A domain	5.5e-32	119.7
900	WD40	WD domain, G-beta repeat	2.7e-07	37.7
901	zf-C2H2	Zinc finger, C2H2 type	4e-156	532.1
903	ras	Ras family	6.6e-101	348.6

SEQ ID NO:	PFAM NAME	DESCRIPTION	p-value	PFAM SCORE
904	Armadillo_seg	Armadillo/beta-catenin-like repeats	1.1e-06	35.6
906	FH2	Formin Homology 2 Domain	4.5e-112	385.7
907	Cytidylyltransf	Cytidylyltransferase	1.4e-05	29.3
908	pkinase	Eukaryotic protein kinase domain	1.2e-64	228.2
909	pkinase	Eukaryotic protein kinase domain	8.5e-70	245.3
910	pkinase	Eukaryotic protein kinase domain	2.9e-42	153.8
911	pkinase	Eukaryotic protein kinase domain	1.2e-35	131.8
912	PHD	PHD-finger	5.1e-06	33.4
913	PHD	PHD-finger	5.5e-16	66.5
916	filament	Intermediate filament proteins	9.7e-121	414.5
917	LIM	LIM domain containing proteins	5.9e-15	57.9
918	SAM	SAM domain (Sterile alpha motif)	4.3e-16	66.9
922	Acylphosphatase	Acylphosphatase	2.9e-63	223.6
924	ig	Immunoglobulin domain	1.3e-08	32.8
925	Acyl-CoA_dh	Acyl-CoA dehydrogenase	2.4e-131	449.8
927	7tm_1	7 transmembrane receptor (rhodopsin family)	2.9e-45	145.9
928	globin	Globin	2.4e-52	186.9
929	sugar_tr	Sugar (and other) transporter	1.2e-16	68.8
932	Collagen	Collagen triple helix repeat (20 copies)	0.00097	9.7
933	HMG_box	HMG (high mobility group) box	7.8e-34	125.8
934	SEA	SEA domain	0.0021	24.7
935	ras	Ras family	6.4e-59	209.2
936	CH	Calponin homology (CH) domain	3.8e-21	83.7
937	voltage_CLC	Voltage gated chloride channels	1.9e-199	676.0
938	homeobox	Homeobox domain	1.9e-25	98.0
940	pkinase	Eukaryotic protein kinase domain	9.9e-58	205.2
942	Myosin_tail	Myosin tail	3.7e-09	38.2
943	zf-C2H2	Zinc finger, C2H2 type	2.2e-92	320.3
945	Clat_adaptor_s	Clathrin adaptor complex small chain	1.3e-76	268.0
946	sugar_tr	Sugar (and other) transporter	0.017	-122.8
947	tRNA-synt_1e	tRNA synthetases class I (C)	0.00097	15.6
948	PHD	PHD-finger	2.2e-17	71.2
951	sugar_tr	Sugar (and other) transporter	0.0082	-113.9
952	mito_carr	Mitochondrial carrier proteins	1.7e-54	189.7
953	myb_DNA-binding	Myb-like DNA-binding domain	4.5e-20	80.1
955	ketoacyl-synt	Beta-ketoacyl synthase	7.1e-133	454.8
957	aldo_ket_red	Aldo/keto reductase family	1.5e-98	340.8
959	Kelch	Kelch motif	0.02	20.8
961	ras	Ras family	2.2e-29	111.1
964	homeobox	Homeobox domain	5.4e-22	86.5
965	PH	PH domain	3e-21	80.9
966	zf-C3HC4	Zinc finger, C3HC4 type (RING finger)	2.2e-09	34.7
967	Ribosomal_L29	Ribosomal L29 protein	1.6e-15	65.0
970	FAD_binding_2	FAD binding domain	8.9e-47	166.6
971	rve	Integrase core domain	0.00015	19.8
972	Glycos_transf_2	Glycosyl transferases	2.1e-21	84.5
974	Ribosomal_L10	Ribosomal protein L10	3.3e-48	173.6
975	7tm_1	7 transmembrane receptor (rhodopsin family)	1.6e-37	121.3
976	zf-C4	Zinc finger, C4 type (two domains)	2.1e-52	178.5
977	zf-C2H2	Zinc finger, C2H2 type	6.6e-150	511.4
978	FTHFS	Formate--tetrahydrofolate ligase	0	1367.2
982	Renal_dipeptase	Renal dipeptidase	1.3e-73	258.0
984	A_deaminase	Adenosine/AMP deaminase	2.6e-05	-48.6

TABLE 5

SEQ ID NO: of full-length nucleotide sequence	SEQ ID NO: of full-length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Priority docket number_corresponding SEQ ID NO: in priority application	SEQ ID NO: in U.S.S.N. 09/496,914
1	985	1969	2953	787CIP2_1	150
2	986	1970	2954	787CIP2_2	223
3	987	1971	2955	787CIP2_3	1884
4	988	1972	2956	787CIP2_4	2123
5	989	1973	2957	787CIP2_5	2313
6	990	1974	2958	787CIP2_6	3284
7	991	1975	2959	787CIP2_7	3324
8	992	1976	2960	787CIP2_8	6182
9	993	1977	2961	787CIP2_9	6210
10	994	1978	2962	787CIP2_10	6213
11	995	1979	2963	787CIP2_11	6257
12	996	1980	2964	787CIP2_12	6294
13	997	1981	2965	787CIP2_13	6294
14	998	1982	2966	787CIP2_14	6330
15	999	1983	2967	787CIP2_15	6364
16	1000	1984	2968	787CIP2_16	6455
17	1001	1985	2969	787CIP2_17	6486
18	1002	1986	2970	787CIP2_18	6503
19	1003	1987	2971	787CIP2_19	6528
20	1004	1988	2972	787CIP2_20	6572
21	1005	1989	2973	787CIP2_21	6578
22	1006	1990	2974	787CIP2_22	6593
23	1007	1991	2975	787CIP2_23	6603
24	1008	1992	2976	787CIP2_24	6603
25	1009	1993	2977	787CIP2_25	6679
26	1010	1994	2978	787CIP2_26	6744
27	1011	1995	2979	787CIP2_27	6762
28	1012	1996	2980	787CIP2_28	6770
29	1013	1997	2981	787CIP2_29	6770
30	1014	1998	2982	787CIP2_30	6787
31	1015	1999	2983	787CIP2_31	6858
32	1016	2000	2984	787CIP2_32	6866
33	1017	2001	2985	787CIP2_33	6938
34	1018	2002	2986	787CIP2_34	6938
35	1019	2003	2987	787CIP2_35	6977
36	1020	2004	2988	787CIP2_36	7001
37	1021	2005	2989	787CIP2_37	7002
38	1022	2006	2990	787CIP2_38	7004
39	1023	2007	2991	787CIP2_39	7005
40	1024	2008	2992	787CIP2_40	7006
41	1025	2009	2993	787CIP2_41	7008
42	1026	2010	2994	787CIP2_42	7014
43	1027	2011	2995	787CIP2_43	7021
44	1028	2012	2996	787CIP2_44	7022
45	1029	2013	2997	787CIP2_46	7057
46	1030	2014	2998	787CIP2_47	7058
47	1031	2015	2999	787CIP2_49	7088
48	1032	2016	3000	787CIP2_50	7089
49	1033	2017	3001	787CIP2_51	7182
50	1034	2018	3002	787CIP2_52	7489
51	1035	2019	3003	787CIP2_53	7564
52	1036	2020	3004	787CIP2_54	7566
53	1037	2021	3005	787CIP2_55	7587

54	1038	2022	3006	787CIP2_56	7591
55	1039	2023	3007	787CIP2_57	7600
56	1040	2024	3008	787CIP2_58	7604
57	1041	2025	3009	787CIP2_59	7612
58	1042	2026	3010	787CIP2_60	7613
59	1043	2027	3011	787CIP2_61	7615
60	1044	2028	3012	787CIP2_62	7616
61	1045	2029	3013	787CIP2_63	7617
62	1046	2030	3014	787CIP2_64	7623
63	1047	2031	3015	787CIP2_65	7625
64	1048	2032	3016	787CIP2_66	7625
65	1049	2033	3017	787CIP2_67	7630
66	1050	2034	3018	787CIP2_68	7638
67	1051	2035	3019	787CIP2_69	7640
68	1052	2036	3020	787CIP2_70	7670
69	1053	2037	3021	787CIP2_71	7676
70	1054	2038	3022	787CIP2_72	7688
71	1055	2039	3023	787CIP2_73	7690
72	1056	2040	3024	787CIP2_74	7700
73	1057	2041	3025	787CIP2_75	7774
74	1058	2042	3026	787CIP2_76	7784
75	1059	2043	3027	787CIP2_77	7785
76	1060	2044	3028	787CIP2_78	7792
77	1061	2045	3029	787CIP2_79	7798
78	1062	2046	3030	787CIP2_80	7807
79	1063	2047	3031	787CIP2_81	7810
80	1064	2048	3032	787CIP2_82	7812
81	1065	2049	3033	787CIP2_83	7816
82	1066	2050	3034	787CIP2_84	7826
83	1067	2051	3035	787CIP2_85	7842
84	1068	2052	3036	787CIP2_86	7850
85	1069	2053	3037	787CIP2_87	7865
86	1070	2054	3038	787CIP2_88	7882
87	1071	2055	3039	787CIP2_89	7891
88	1072	2056	3040	787CIP2_90	7892
89	1073	2057	3041	787CIP2_91	7896
90	1074	2058	3042	787CIP2_92	7896
91	1075	2059	3043	787CIP2_93	7907
92	1076	2060	3044	787CIP2_94	7913
93	1077	2061	3045	787CIP2_95	7914
94	1078	2062	3046	787CIP2_96	7915
95	1079	2063	3047	787CIP2_97	7920
96	1080	2064	3048	787CIP2_98	7921
97	1081	2065	3049	787CIP2_99	7924
98	1082	2066	3050	787CIP2_100	7927
99	1083	2067	3051	787CIP2_101	7929
100	1084	2068	3052	787CIP2_102	7937
101	1085	2069	3053	787CIP2_103	7940
102	1086	2070	3054	787CIP2_104	7942
103	1087	2071	3055	787CIP2_105	7944
104	1088	2072	3056	787CIP2_106	7951
105	1089	2073	3057	787CIP2_107	7951
106	1090	2074	3058	787CIP2_108	7962
107	1091	2075	3059	787CIP2_109	7964
108	1092	2076	3060	787CIP2_110	7977
109	1093	2077	3061	787CIP2_111	7978
110	1094	2078	3062	787CIP2_112	7980
111	1095	2079	3063	787CIP2_113	7982
112	1096	2080	3064	787CIP2_114	8000
113	1097	2081	3065	787CIP2_115	8003

114	1098	2082	3066	787CIP2_116	8004
115	1099	2083	3067	787CIP2_117	8007
116	1100	2084	3068	787CIP2_118	8008
117	1101	2085	3069	787CIP2_119	8009
118	1102	2086	3070	787CIP2_120	8013
119	1103	2087	3071	787CIP2_121	8017
120	1104	2088	3072	787CIP2_122	8018
121	1105	2089	3073	787CIP2_123	8021
122	1106	2090	3074	787CIP2_124	8022
123	1107	2091	3075	787CIP2_125	8023
124	1108	2092	3076	787CIP2_126	8023
125	1109	2093	3077	787CIP2_127	8024
126	1110	2094	3078	787CIP2_128	8026
127	1111	2095	3079	787CIP2_129	8028
128	1112	2096	3080	787CIP2_130	8036
129	1113	2097	3081	787CIP2_131	8038
130	1114	2098	3082	787CIP2_132	8045
131	1115	2099	3083	787CIP2_133	8045
132	1116	2100	3084	787CIP2_134	8048
133	1117	2101	3085	787CIP2_135	8048
134	1118	2102	3086	787CIP2_136	8052
135	1119	2103	3087	787CIP2_137	8053
136	1120	2104	3088	787CIP2_138	8055
137	1121	2105	3089	787CIP2_139	8059
138	1122	2106	3090	787CIP2_140	8061
139	1123	2107	3091	787CIP2_141	8062
140	1124	2108	3092	787CIP2_142	8063
141	1125	2109	3093	787CIP2_143	8064
142	1126	2110	3094	787CIP2_144	8065
143	1127	2111	3095	787CIP2_145	8068
144	1128	2112	3096	787CIP2_146	8069
145	1129	2113	3097	787CIP2_147	8070
146	1130	2114	3098	787CIP2_148	8074
147	1131	2115	3099	787CIP2_149	8076
148	1132	2116	3100	787CIP2_150	8077
149	1133	2117	3101	787CIP2_151	8078
150	1134	2118	3102	787CIP2_152	8079
151	1135	2119	3103	787CIP2_153	8087
152	1136	2120	3104	787CIP2_154	8091
153	1137	2121	3105	787CIP2_155	8100
154	1138	2122	3106	787CIP2_156	8105
155	1139	2123	3107	787CIP2_157	8106
156	1140	2124	3108	787CIP2_158	8108
157	1141	2125	3109	787CIP2_159	8109
158	1142	2126	3110	787CIP2_160	8110
159	1143	2127	3111	787CIP2_161	8112
160	1144	2128	3112	787CIP2_162	8116
161	1145	2129	3113	787CIP2_163	8118
162	1146	2130	3114	787CIP2_164	8124
163	1147	2131	3115	787CIP2_165	8125
164	1148	2132	3116	787CIP2_166	8127
165	1149	2133	3117	787CIP2_167	8132
166	1150	2134	3118	787CIP2_168	8135
167	1151	2135	3119	787CIP2_169	8137
168	1152	2136	3120	787CIP2_170	8139
169	1153	2137	3121	787CIP2_171	8140
170	1154	2138	3122	787CIP2_172	8140
171	1155	2139	3123	787CIP2_173	8140
172	1156	2140	3124	787CIP2_174	8141
173	1157	2141	3125	787CIP2_175	8147

174	1158	2142	3126	787CIP2_176	8149
175	1159	2143	3127	787CIP2_177	8150
176	1160	2144	3128	787CIP2_178	8157
177	1161	2145	3129	787CIP2_179	8161
178	1162	2146	3130	787CIP2_180	8162
179	1163	2147	3131	787CIP2_181	8165
180	1164	2148	3132	787CIP2_182	8166
181	1165	2149	3133	787CIP2_183	8167
182	1166	2150	3134	787CIP2_184	8169
183	1167	2151	3135	787CIP2_185	8170
184	1168	2152	3136	787CIP2_186	8172
185	1169	2153	3137	787CIP2_187	8173
186	1170	2154	3138	787CIP2_188	8174
187	1171	2155	3139	787CIP2_189	8174
188	1172	2156	3140	787CIP2_191	8182
189	1173	2157	3141	787CIP2_192	8186
190	1174	2158	3142	787CIP2_193	8188
191	1175	2159	3143	787CIP2_194	8191
192	1176	2160	3144	787CIP2_195	8192
193	1177	2161	3145	787CIP2_196	8193
194	1178	2162	3146	787CIP2_197	8194
195	1179	2163	3147	787CIP2_198	8195
196	1180	2164	3148	787CIP2_199	8196
197	1181	2165	3149	787CIP2_200	8200
198	1182	2166	3150	787CIP2_201	8201
199	1183	2167	3151	787CIP2_202	8202
200	1184	2168	3152	787CIP2_203	8205
201	1185	2169	3153	787CIP2_204	8206
202	1186	2170	3154	787CIP2_205	8207
203	1187	2171	3155	787CIP2_206	8208
204	1188	2172	3156	787CIP2_207	8209
205	1189	2173	3157	787CIP2_208	8210
206	1190	2174	3158	787CIP2_209	8211
207	1191	2175	3159	787CIP2_210	8212
208	1192	2176	3160	787CIP2_211	8213
209	1193	2177	3161	787CIP2_212	8214
210	1194	2178	3162	787CIP2_213	8215
211	1195	2179	3163	787CIP2_214	8216
212	1196	2180	3164	787CIP2_215	8217
213	1197	2181	3165	787CIP2_217	8221
214	1198	2182	3166	787CIP2_218	8222
215	1199	2183	3167	787CIP2_219	8223
216	1200	2184	3168	787CIP2_220	8224
217	1201	2185	3169	787CIP2_221	8225
218	1202	2186	3170	787CIP2_222	8227
219	1203	2187	3171	787CIP2_223	8232
220	1204	2188	3172	787CIP2_224	8235
221	1205	2189	3173	787CIP2_225	8236
222	1206	2190	3174	787CIP2_227	8238
223	1207	2191	3175	787CIP2_228	8239
224	1208	2192	3176	787CIP2_229	8240
225	1209	2193	3177	787CIP2_230	8242
226	1210	2194	3178	787CIP2_231	8246
227	1211	2195	3179	787CIP2_232	8252
228	1212	2196	3180	787CIP2_233	8257
229	1213	2197	3181	787CIP2_234	8288
230	1214	2198	3182	787CIP2_235	8310
231	1215	2199	3183	787CIP2_236	8311
232	1216	2200	3184	787CIP2_237	8315
233	1217	2201	3185	787CIP2_238	8318

234	1218	2202	3186	787CIP2_239	8326
235	1219	2203	3187	787CIP2_240	8326
236	1220	2204	3188	787CIP2_241	8336
237	1221	2205	3189	787CIP2_242	8351
238	1222	2206	3190	787CIP2_243	8364
239	1223	2207	3191	787CIP2_244	8372
240	1224	2208	3192	787CIP2_245	8376
241	1225	2209	3193	787CIP2_246	8377
242	1226	2210	3194	787CIP2_247	8382
243	1227	2211	3195	787CIP2_248	8404
244	1228	2212	3196	787CIP2_249	8410
245	1229	2213	3197	787CIP2_250	8419
246	1230	2214	3198	787CIP2_251	8430
247	1231	2215	3199	787CIP2_252	8448
248	1232	2216	3200	787CIP2_253	8458
249	1233	2217	3201	787CIP2_254	8461
250	1234	2218	3202	787CIP2_255	8466
251	1235	2219	3203	787CIP2_256	8468
252	1236	2220	3204	787CIP2_257	8477
253	1237	2221	3205	787CIP2_258	8481
254	1238	2222	3206	787CIP2_259	8491
255	1239	2223	3207	787CIP2_260	8503
256	1240	2224	3208	787CIP2_261	8513
257	1241	2225	3209	787CIP2_262	8514
258	1242	2226	3210	787CIP2_263	8518
259	1243	2227	3211	787CIP2_264	8547
260	1244	2228	3212	787CIP2_265	8549
261	1245	2229	3213	787CIP2_266	8549
262	1246	2230	3214	787CIP2_267	8549
263	1247	2231	3215	787CIP2_268	8550
264	1248	2232	3216	787CIP2_269	8603
265	1249	2233	3217	787CIP2_270	8625
266	1250	2234	3218	787CIP2_271	8625
267	1251	2235	3219	787CIP2_272	8633
268	1252	2236	3220	787CIP2_273	8648
269	1253	2237	3221	787CIP2_274	8654
270	1254	2238	3222	787CIP2_275	8671
271	1255	2239	3223	787CIP2_276	8733
272	1256	2240	3224	787CIP2_277	8735
273	1257	2241	3225	787CIP2_278	8747
274	1258	2242	3226	787CIP2_279	8748
275	1259	2243	3227	787CIP2_280	8753
276	1260	2244	3228	787CIP2_281	8770
277	1261	2245	3229	787CIP2_282	8777
278	1262	2246	3230	787CIP2_283	8828
279	1263	2247	3231	787CIP2_284	8836
280	1264	2248	3232	787CIP2_285	8842
281	1265	2249	3233	787CIP2_286	8842
282	1266	2250	3234	787CIP2_287	8850
283	1267	2251	3235	787CIP2_288	8851
284	1268	2252	3236	787CIP2_289	8852
285	1269	2253	3237	787CIP2_290	8853
286	1270	2254	3238	787CIP2_291	8854
287	1271	2255	3239	787CIP2_292	9084
288	1272	2256	3240	787CIP2_293	9099
289	1273	2257	3241	787CIP2_294	9691
290	1274	2258	3242	787CIP2_295	9699
291	1275	2259	3243	787CIP2_296	9883
292	1276	2260	3244	787CIP2_297	9886
293	1277	2261	3245	787CIP2_298	10334

294	1278	2262	3246	787CIP2_299	10335
295	1279	2263	3247	787CIP2_300	10336
296	1280	2264	3248	787CIP2_301	10338
297	1281	2265	3249	787CIP2_302	10339
298	1282	2266	3250	787CIP2_304	10342
299	1283	2267	3251	787CIP2_305	10342
300	1284	2268	3252	787CIP2_306	10343
301	1285	2269	3253	787CIP2_307	10344
302	1286	2270	3254	787CIP2_308	10345
303	1287	2271	3255	787CIP2_309	10346
304	1288	2272	3256	787CIP2_310	10347
305	1289	2273	3257	787CIP2_311	10348
306	1290	2274	3258	787CIP2_312	10349
307	1291	2275	3259	787CIP2_314	10351
308	1292	2276	3260	787CIP2_315	10352
309	1293	2277	3261	787CIP2_316	10353
310	1294	2278	3262	787CIP2_317	10354
311	1295	2279	3263	787CIP2_318	10355
312	1296	2280	3264	787CIP2_319	10356
313	1297	2281	3265	787CIP2_320	10357
314	1298	2282	3266	787CIP2_321	10358
315	1299	2283	3267	787CIP2_322	10360
316	1300	2284	3268	787CIP2_323	10361
317	1301	2285	3269	787CIP2_324	10362
318	1302	2286	3270	787CIP2_325	10363
319	1303	2287	3271	787CIP2_326	10365
320	1304	2288	3272	787CIP2_327	10366
321	1305	2289	3273	787CIP2_328	10367
322	1306	2290	3274	787CIP2_329	10369
323	1307	2291	3275	787CIP2_330	10370
324	1308	2292	3276	787CIP2_331	10371
325	1309	2293	3277	787CIP2_332	10372
326	1310	2294	3278	787CIP2_333	10373
327	1311	2295	3279	787CIP2_334	10375
328	1312	2296	3280	787CIP2_335	10377
329	1313	2297	3281	787CIP2_336	10379
330	1314	2298	3282	787CIP2_337	10381
331	1315	2299	3283	787CIP2_338	10382
332	1316	2300	3284	787CIP2_339	10383
333	1317	2301	3285	787CIP2_340	10384
334	1318	2302	3286	787CIP2_341	10385
335	1319	2303	3287	787CIP2_342	10386
336	1320	2304	3288	787CIP2_343	10387
337	1321	2305	3289	787CIP2_346	10391
338	1322	2306	3290	787CIP2_348	10393
339	1323	2307	3291	787CIP2_349	10394
340	1324	2308	3292	787CIP2_350	10395
341	1325	2309	3293	787CIP2_351	10396
342	1326	2310	3294	787CIP2_352	10397
343	1327	2311	3295	787CIP2_353	10399
344	1328	2312	3296	787CIP2_354	10400
345	1329	2313	3297	787CIP2_355	10401
346	1330	2314	3298	787CIP2_357	10403
347	1331	2315	3299	787CIP2_358	10404
348	1332	2316	3300	787CIP2_359	10407
349	1333	2317	3301	787CIP2_360	10408
350	1334	2318	3302	787CIP2_361	10409
351	1335	2319	3303	787CIP2_362	10410
352	1336	2320	3304	787CIP2B_1	44
353	1337	2321	3305	787CIP2B_2	50

354	1338	2322	3306	787CIP2B_3	93
355	1339	2323	3307	787CIP2B_4	224
356	1340	2324	3308	787CIP2B_5	318
357	1341	2325	3309	787CIP2B_6	318
358	1342	2326	3310	787CIP2B_7	795
359	1343	2327	3311	787CIP2B_8	857
360	1344	2328	3312	787CIP2B_9	924
361	1345	2329	3313	787CIP2B_10	944
362	1346	2330	3314	787CIP2B_11	944
363	1347	2331	3315	787CIP2B_12	967
364	1348	2332	3316	787CIP2B_13	1055
365	1349	2333	3317	787CIP2B_14	1091
366	1350	2334	3318	787CIP2B_15	1225
367	1351	2335	3319	787CIP2B_16	1257
368	1352	2336	3320	787CIP2B_17	1289
369	1353	2337	3321	787CIP2B_18	1292
370	1354	2338	3322	787CIP2B_19	1455
371	1355	2339	3323	787CIP2B_20	1488
372	1356	2340	3324	787CIP2B_21	1666
373	1357	2341	3325	787CIP2B_22	1811
374	1358	2342	3326	787CIP2B_23	1885
375	1359	2343	3327	787CIP2B_24	1911
376	1360	2344	3328	787CIP2B_25	1935
377	1361	2345	3329	787CIP2B_26	1971
378	1362	2346	3330	787CIP2B_27	1989
379	1363	2347	3331	787CIP2B_28	2041
380	1364	2348	3332	787CIP2B_29	2178
381	1365	2349	3333	787CIP2B_30	2237
382	1366	2350	3334	787CIP2B_31	2279
383	1367	2351	3335	787CIP2B_32	2338
384	1368	2352	3336	787CIP2B_33	2351
385	1369	2353	3337	787CIP2B_34	2405
386	1370	2354	3338	787CIP2B_35	2531
387	1371	2355	3339	787CIP2B_36	2584
388	1372	2356	3340	787CIP2B_37	2608
389	1373	2357	3341	787CIP2B_38	2655
390	1374	2358	3342	787CIP2B_39	2656
391	1375	2359	3343	787CIP2B_40	2866
392	1376	2360	3344	787CIP2B_41	3015
393	1377	2361	3345	787CIP2B_42	3015
394	1378	2362	3346	787CIP2B_43	3043
395	1379	2363	3347	787CIP2B_44	3986
396	1380	2364	3348	787CIP2B_45	4647
397	1381	2365	3349	787CIP2B_46	4659
398	1382	2366	3350	787CIP2B_47	5032
399	1383	2367	3351	787CIP2B_48	5244
400	1384	2368	3352	787CIP2B_49	5268
401	1385	2369	3353	787CIP2B_50	5281
402	1386	2370	3354	787CIP2B_51	5282
403	1387	2371	3355	787CIP2B_52	6147
404	1388	2372	3356	787CIP2B_53	6178
405	1389	2373	3357	787CIP2B_54	6184
406	1390	2374	3358	787CIP2B_55	6187
407	1391	2375	3359	787CIP2B_56	6190
408	1392	2376	3360	787CIP2B_57	6191
409	1393	2377	3361	787CIP2B_58	6194
410	1394	2378	3362	787CIP2B_59	6196
411	1395	2379	3363	787CIP2B_60	6201
412	1396	2380	3364	787CIP2B_61	6208
413	1397	2381	3365	787CIP2B_62	6214

414	1398	2382	3366	787CIP2B_63	6217
415	1399	2383	3367	787CIP2B_64	6220
416	1400	2384	3368	787CIP2B_65	6221
417	1401	2385	3369	787CIP2B_66	6222
418	1402	2386	3370	787CIP2B_67	6223
419	1403	2387	3371	787CIP2B_68	6223
420	1404	2388	3372	787CIP2B_69	6226
421	1405	2389	3373	787CIP2B_70	6227
422	1406	2390	3374	787CIP2B_71	6229
423	1407	2391	3375	787CIP2B_72	6248
424	1408	2392	3376	787CIP2B_73	6260
425	1409	2393	3377	787CIP2B_74	6264
426	1410	2394	3378	787CIP2B_75	6269
427	1411	2395	3379	787CIP2B_76	6269
428	1412	2396	3380	787CIP2B_77	6275
429	1413	2397	3381	787CIP2B_78	6276
430	1414	2398	3382	787CIP2B_79	6280
431	1415	2399	3383	787CIP2B_80	6287
432	1416	2400	3384	787CIP2B_81	6290
433	1417	2401	3385	787CIP2B_82	6293
434	1418	2402	3386	787CIP2B_83	6305
435	1419	2403	3387	787CIP2B_84	6308
436	1420	2404	3388	787CIP2B_85	6309
437	1421	2405	3389	787CIP2B_86	6312
438	1422	2406	3390	787CIP2B_87	6314
439	1423	2407	3391	787CIP2B_88	6316
440	1424	2408	3392	787CIP2B_89	6336
441	1425	2409	3393	787CIP2B_90	6341
442	1426	2410	3394	787CIP2B_91	6343
443	1427	2411	3395	787CIP2B_92	6346
444	1428	2412	3396	787CIP2B_93	6357
445	1429	2413	3397	787CIP2B_94	6359
446	1430	2414	3398	787CIP2B_95	6367
447	1431	2415	3399	787CIP2B_96	6383
448	1432	2416	3400	787CIP2B_97	6385
449	1433	2417	3401	787CIP2B_98	6396
450	1434	2418	3402	787CIP2B_99	6396
451	1435	2419	3403	787CIP2B_100	6403
452	1436	2420	3404	787CIP2B_101	6405
453	1437	2421	3405	787CIP2B_102	6414
454	1438	2422	3406	787CIP2B_103	6418
455	1439	2423	3407	787CIP2B_104	6422
456	1440	2424	3408	787CIP2B_105	6425
457	1441	2425	3409	787CIP2B_106	6436
458	1442	2426	3410	787CIP2B_107	6471
459	1443	2427	3411	787CIP2B_108	6474
460	1444	2428	3412	787CIP2B_109	6482
461	1445	2429	3413	787CIP2B_110	6504
462	1446	2430	3414	787CIP2B_111	6510
463	1447	2431	3415	787CIP2B_112	6515
464	1448	2432	3416	787CIP2B_113	6529
465	1449	2433	3417	787CIP2B_114	6535
466	1450	2434	3418	787CIP2B_115	6536
467	1451	2435	3419	787CIP2B_116	6536
468	1452	2436	3420	787CIP2B_117	6541
469	1453	2437	3421	787CIP2B_118	6542
470	1454	2438	3422	787CIP2B_119	6547
471	1455	2439	3423	787CIP2B_120	6548
472	1456	2440	3424	787CIP2B_121	6552
473	1457	2441	3425	787CIP2B_122	6552

474	1458	2442	3426	787CIP2B_123	6555
475	1459	2443	3427	787CIP2B_124	6560
476	1460	2444	3428	787CIP2B_125	6566
477	1461	2445	3429	787CIP2B_126	6576
478	1462	2446	3430	787CIP2B_127	6584
479	1463	2447	3431	787CIP2B_128	6588
480	1464	2448	3432	787CIP2B_129	6589
481	1465	2449	3433	787CIP2B_130	6590
482	1466	2450	3434	787CIP2B_131	6597
483	1467	2451	3435	787CIP2B_132	6600
484	1468	2452	3436	787CIP2B_133	6602
485	1469	2453	3437	787CIP2B_134	6604
486	1470	2454	3438	787CIP2B_135	6605
487	1471	2455	3439	787CIP2B_136	6608
488	1472	2456	3440	787CIP2B_137	6610
489	1473	2457	3441	787CIP2B_138	6614
490	1474	2458	3442	787CIP2B_139	6623
491	1475	2459	3443	787CIP2B_140	6629
492	1476	2460	3444	787CIP2B_141	6631
493	1477	2461	3445	787CIP2B_142	6631
494	1478	2462	3446	787CIP2B_143	6631
495	1479	2463	3447	787CIP2B_144	6632
496	1480	2464	3448	787CIP2B_145	6633
497	1481	2465	3449	787CIP2B_146	6634
498	1482	2466	3450	787CIP2B_147	6635
499	1483	2467	3451	787CIP2B_148	6639
500	1484	2468	3452	787CIP2B_149	6649
501	1485	2469	3453	787CIP2B_150	6651
502	1486	2470	3454	787CIP2B_151	6655
503	1487	2471	3455	787CIP2B_152	6658
504	1488	2472	3456	787CIP2B_153	6667
505	1489	2473	3457	787CIP2B_154	6672
506	1490	2474	3458	787CIP2B_155	6682
507	1491	2475	3459	787CIP2B_156	6683
508	1492	2476	3460	787CIP2B_157	6687
509	1493	2477	3461	787CIP2B_158	6687
510	1494	2478	3462	787CIP2B_159	6688
511	1495	2479	3463	787CIP2B_160	6696
512	1496	2480	3464	787CIP2B_161	6701
513	1497	2481	3465	787CIP2B_162	6707
514	1498	2482	3466	787CIP2B_163	6712
515	1499	2483	3467	787CIP2B_164	6714
516	1500	2484	3468	787CIP2B_165	6720
517	1501	2485	3469	787CIP2B_166	6721
518	1502	2486	3470	787CIP2B_167	6722
519	1503	2487	3471	787CIP2B_168	6736
520	1504	2488	3472	787CIP2B_169	6740
521	1505	2489	3473	787CIP2B_170	6740
522	1506	2490	3474	787CIP2B_171	6760
523	1507	2491	3475	787CIP2B_172	6775
524	1508	2492	3476	787CIP2B_173	6784
525	1509	2493	3477	787CIP2B_174	6793
526	1510	2494	3478	787CIP2B_175	6795
527	1511	2495	3479	787CIP2B_176	6796
528	1512	2496	3480	787CIP2B_177	6807
529	1513	2497	3481	787CIP2B_178	6808
530	1514	2498	3482	787CIP2B_179	6810
531	1515	2499	3483	787CIP2B_180	6815
532	1516	2500	3484	787CIP2B_181	6819
533	1517	2501	3485	787CIP2B_182	6821

534	1518	2502	3486	787CIP2B_183	6827
535	1519	2503	3487	787CIP2B_184	6829
536	1520	2504	3488	787CIP2B_185	6830
537	1521	2505	3489	787CIP2B_186	6835
538	1522	2506	3490	787CIP2B_187	6848
539	1523	2507	3491	787CIP2B_188	6849
540	1524	2508	3492	787CIP2B_189	6851
541	1525	2509	3493	787CIP2B_190	6851
542	1526	2510	3494	787CIP2B_191	6863
543	1527	2511	3495	787CIP2B_192	6869
544	1528	2512	3496	787CIP2B_193	6874
545	1529	2513	3497	787CIP2B_194	6887
546	1530	2514	3498	787CIP2B_195	6890
547	1531	2515	3499	787CIP2B_196	6894
548	1532	2516	3500	787CIP2B_197	6899
549	1533	2517	3501	787CIP2B_198	6900
550	1534	2518	3502	787CIP2B_199	6903
551	1535	2519	3503	787CIP2B_200	6910
552	1536	2520	3504	787CIP2B_201	6913
553	1537	2521	3505	787CIP2B_202	6918
554	1538	2522	3506	787CIP2B_203	6923
555	1539	2523	3507	787CIP2B_204	6926
556	1540	2524	3508	787CIP2B_205	6929
557	1541	2525	3509	787CIP2B_206	6929
558	1542	2526	3510	787CIP2B_207	6932
559	1543	2527	3511	787CIP2B_208	6941
560	1544	2528	3512	787CIP2B_209	6951
561	1545	2529	3513	787CIP2B_210	6954
562	1546	2530	3514	787CIP2B_211	6954
563	1547	2531	3515	787CIP2B_212	6956
564	1548	2532	3516	787CIP2B_213	6957
565	1549	2533	3517	787CIP2B_214	6960
566	1550	2534	3518	787CIP2B_215	6966
567	1551	2535	3519	787CIP2B_216	6968
568	1552	2536	3520	787CIP2B_217	6969
569	1553	2537	3521	787CIP2B_218	6970
570	1554	2538	3522	787CIP2B_219	6971
571	1555	2539	3523	787CIP2B_220	6989
572	1556	2540	3524	787CIP2B_221	6990
573	1557	2541	3525	787CIP2B_223	6996
574	1558	2542	3526	787CIP2B_224	6997
575	1559	2543	3527	787CIP2B_225	7009
576	1560	2544	3528	787CIP2B_226	7016
577	1561	2545	3529	787CIP2B_227	7023
578	1562	2546	3530	787CIP2B_228	7023
579	1563	2547	3531	787CIP2B_229	7035
580	1564	2548	3532	787CIP2B_230	7038
581	1565	2549	3533	787CIP2B_231	7039
582	1566	2550	3534	787CIP2B_232	7040
583	1567	2551	3535	787CIP2B_233	7041
584	1568	2552	3536	787CIP2B_234	7044
585	1569	2553	3537	787CIP2B_235	7059
586	1570	2554	3538	787CIP2B_236	7060
587	1571	2555	3539	787CIP2B_237	7063
588	1572	2556	3540	787CIP2B_238	7067
589	1573	2557	3541	787CIP2B_239	7070
590	1574	2558	3542	787CIP2B_240	7071
591	1575	2559	3543	787CIP2B_241	7079
592	1576	2560	3544	787CIP2B_242	7085
593	1577	2561	3545	787CIP2B_243	7148

594	1578	2562	3546	787CIP2B_244	7156
595	1579	2563	3547	787CIP2B_245	7156
596	1580	2564	3548	787CIP2B_246	7171
597	1581	2565	3549	787CIP2B_248	7265
598	1582	2566	3550	787CIP2B_249	7268
599	1583	2567	3551	787CIP2B_250	7308
600	1584	2568	3552	787CIP2B_251	7336
601	1585	2569	3553	787CIP2B_252	7347
602	1586	2570	3554	787CIP2B_253	7405
603	1587	2571	3555	787CIP2B_254	7405
604	1588	2572	3556	787CIP2B_255	7412
605	1589	2573	3557	787CIP2B_256	7412
606	1590	2574	3558	787CIP2B_257	7436
607	1591	2575	3559	787CIP2B_258	7436
608	1592	2576	3560	787CIP2B_259	7454
609	1593	2577	3561	787CIP2B_260	7476
610	1594	2578	3562	787CIP2B_261	7598
611	1595	2579	3563	787CIP2B_262	7619
612	1596	2580	3564	787CIP2B_263	7644
613	1597	2581	3565	787CIP2B_264	7648
614	1598	2582	3566	787CIP2B_265	7659
615	1599	2583	3567	787CIP2B_266	7661
616	1600	2584	3568	787CIP2B_267	7669
617	1601	2585	3569	787CIP2B_268	7686
618	1602	2586	3570	787CIP2B_269	7686
619	1603	2587	3571	787CIP2B_270	7694
620	1604	2588	3572	787CIP2B_271	7697
621	1605	2589	3573	787CIP2B_272	7733
622	1606	2590	3574	787CIP2B_273	7734
623	1607	2591	3575	787CIP2B_274	7744
624	1608	2592	3576	787CIP2B_275	7751
625	1609	2593	3577	787CIP2B_276	7756
626	1610	2594	3578	787CIP2B_277	7761
627	1611	2595	3579	787CIP2B_278	7761
628	1612	2596	3580	787CIP2B_279	7776
629	1613	2597	3581	787CIP2B_280	7783
630	1614	2598	3582	787CIP2B_281	7800
631	1615	2599	3583	787CIP2B_282	7800
632	1616	2600	3584	787CIP2B_283	7801
633	1617	2601	3585	787CIP2B_284	7811
634	1618	2602	3586	787CIP2B_285	7817
635	1619	2603	3587	787CIP2B_286	7821
636	1620	2604	3588	787CIP2B_287	7822
637	1621	2605	3589	787CIP2B_288	7841
638	1622	2606	3590	787CIP2B_289	7847
639	1623	2607	3591	787CIP2B_290	7880
640	1624	2608	3592	787CIP2B_291	7910
641	1625	2609	3593	787CIP2B_293	7936
642	1626	2610	3594	787CIP2B_294	7945
643	1627	2611	3595	787CIP2B_295	7948
644	1628	2612	3596	787CIP2B_296	7963
645	1629	2613	3597	787CIP2B_297	7984
646	1630	2614	3598	787CIP2B_298	7985
647	1631	2615	3599	787CIP2B_299	8014
648	1632	2616	3600	787CIP2B_301	8029
649	1633	2617	3601	787CIP2B_302	8043
650	1634	2618	3602	787CIP2B_303	8164
651	1635	2619	3603	787CIP2B_304	8175
652	1636	2620	3604	787CIP2B_305	8250
653	1637	2621	3605	787CIP2B_306	8253

654	1638	2622	3606	787CIP2B_307	8255
655	1639	2623	3607	787CIP2B_308	8258
656	1640	2624	3608	787CIP2B_309	8270
657	1641	2625	3609	787CIP2B_310	8271
658	1642	2626	3610	787CIP2B_311	8272
659	1643	2627	3611	787CIP2B_312	8279
660	1644	2628	3612	787CIP2B_313	8284
661	1645	2629	3613	787CIP2B_314	8285
662	1646	2630	3614	787CIP2B_315	8304
663	1647	2631	3615	787CIP2B_316	8309
664	1648	2632	3616	787CIP2B_317	8320
665	1649	2633	3617	787CIP2B_318	8331
666	1650	2634	3618	787CIP2B_319	8332
667	1651	2635	3619	787CIP2B_320	8332
668	1652	2636	3620	787CIP2B_321	8335
669	1653	2637	3621	787CIP2B_322	8337
670	1654	2638	3622	787CIP2B_323	8353
671	1655	2639	3623	787CIP2B_324	8355
672	1656	2640	3624	787CIP2B_325	8358
673	1657	2641	3625	787CIP2B_326	8361
674	1658	2642	3626	787CIP2B_327	8369
675	1659	2643	3627	787CIP2B_328	8385
676	1660	2644	3628	787CIP2B_329	8397
677	1661	2645	3629	787CIP2B_330	8414
678	1662	2646	3630	787CIP2B_331	8431
679	1663	2647	3631	787CIP2B_332	8433
680	1664	2648	3632	787CIP2B_333	8444
681	1665	2649	3633	787CIP2B_334	8446
682	1666	2650	3634	787CIP2B_335	8460
683	1667	2651	3635	787CIP2B_336	8478
684	1668	2652	3636	787CIP2B_337	8490
685	1669	2653	3637	787CIP2B_338	8505
686	1670	2654	3638	787CIP2B_339	8523
687	1671	2655	3639	787CIP2B_340	8530
688	1672	2656	3640	787CIP2B_341	8533
689	1673	2657	3641	787CIP2B_342	8534
690	1674	2658	3642	787CIP2B_343	8536
691	1675	2659	3643	787CIP2B_344	8537
692	1676	2660	3644	787CIP2B_345	8543
693	1677	2661	3645	787CIP2B_346	8546
694	1678	2662	3646	787CIP2B_347	8553
695	1679	2663	3647	787CIP2B_348	8556
696	1680	2664	3648	787CIP2B_349	8561
697	1681	2665	3649	787CIP2B_350	8562
698	1682	2666	3650	787CIP2B_351	8569
699	1683	2667	3651	787CIP2B_352	8587
700	1684	2668	3652	787CIP2B_353	8597
701	1685	2669	3653	787CIP2B_354	8610
702	1686	2670	3654	787CIP2B_355	8610
703	1687	2671	3655	787CIP2B_356	8615
704	1688	2672	3656	787CIP2B_357	8622
705	1689	2673	3657	787CIP2B_358	8626
706	1690	2674	3658	787CIP2B_359	8628
707	1691	2675	3659	787CIP2B_360	8629
708	1692	2676	3660	787CIP2B_361	8630
709	1693	2677	3661	787CIP2B_362	8632
710	1694	2678	3662	787CIP2B_363	8634
711	1695	2679	3663	787CIP2B_364	8643
712	1696	2680	3664	787CIP2B_365	8644
713	1697	2681	3665	787CIP2B_366	8645

714	1698	2682	3666	787CIP2B_367	8646
715	1699	2683	3667	787CIP2B_368	8657
716	1700	2684	3668	787CIP2B_369	8661
717	1701	2685	3669	787CIP2B_370	8670
718	1702	2686	3670	787CIP2B_371	8692
719	1703	2687	3671	787CIP2B_372	8698
720	1704	2688	3672	787CIP2B_373	8762
721	1705	2689	3673	787CIP2B_374	8768
722	1706	2690	3674	787CIP2B_375	8768
723	1707	2691	3675	787CIP2B_376	8799
724	1708	2692	3676	787CIP2B_377	8806
725	1709	2693	3677	787CIP2B_378	8809
726	1710	2694	3678	787CIP2B_379	8814
727	1711	2695	3679	787CIP2B_380	8822
728	1712	2696	3680	787CIP2B_381	8833
729	1713	2697	3681	787CIP2B_382	8835
730	1714	2698	3682	787CIP2B_383	8877
731	1715	2699	3683	787CIP2B_384	8886
732	1716	2700	3684	787CIP2B_385	9003
733	1717	2701	3685	787CIP2B_386	9157
734	1718	2702	3686	787CIP2B_387	9175
735	1719	2703	3687	787CIP2B_388	9205
736	1720	2704	3688	787CIP2B_389	9260
737	1721	2705	3689	787CIP2B_390	9295
738	1722	2706	3690	787CIP2B_391	9307
739	1723	2707	3691	787CIP2B_392	9307
740	1724	2708	3692	787CIP2B_393	9312
741	1725	2709	3693	787CIP2B_394	9347
742	1726	2710	3694	787CIP2B_395	9370
743	1727	2711	3695	787CIP2B_396	9370
744	1728	2712	3696	787CIP2B_397	9382
745	1729	2713	3697	787CIP2B_398	9591
746	1730	2714	3698	787CIP2B_399	9650
747	1731	2715	3699	787CIP2B_400	9655
748	1732	2716	3700	787CIP2B_401	9663
749	1733	2717	3701	787CIP2B_402	9715
750	1734	2718	3702	787CIP2B_403	9755
751	1735	2719	3703	787CIP2B_404	9766
752	1736	2720	3704	787CIP2B_405	9771
753	1737	2721	3705	787CIP2B_406	9784
754	1738	2722	3706	787CIP2B_407	9925
755	1739	2723	3707	787CIP2B_408	9970
756	1740	2724	3708	787CIP2B_409	9997
757	1741	2725	3709	787CIP2B_410	10008
758	1742	2726	3710	787CIP2B_411	10010
759	1743	2727	3711	787CIP2B_412	10023
760	1744	2728	3712	787CIP2B_413	10043
761	1745	2729	3713	787CIP2B_414	10093
762	1746	2730	3714	787CIP2B_415	10172
763	1747	2731	3715	787CIP2B_416	10184
764	1748	2732	3716	787CIP2B_417	10205
765	1749	2733	3717	787CIP2B_418	10246
766	1750	2734	3718	787CIP2B_419	10298
767	1751	2735	3719	787CIP2C_1	886
768	1752	2736	3720	787CIP2C_2	1028
769	1753	2737	3721	787CIP2C_3	1916
770	1754	2738	3722	787CIP2C_4	2072
771	1755	2739	3723	787CIP2C_5	2424
772	1756	2740	3724	787CIP2C_6	2474
773	1757	2741	3725	787CIP2C_7	2474

774	1758	2742	3726	787CIP2C_8	2887
775	1759	2743	3727	787CIP2C_9	3001
776	1760	2744	3728	787CIP2C_10	3182
777	1761	2745	3729	787CIP2C_11	3182
778	1762	2746	3730	787CIP2C_12	3182
779	1763	2747	3731	787CIP2C_13	3193
780	1764	2748	3732	787CIP2C_14	3196
781	1765	2749	3733	787CIP2C_15	3224
782	1766	2750	3734	787CIP2C_16	3225
783	1767	2751	3735	787CIP2C_17	3234
784	1768	2752	3736	787CIP2C_18	3241
785	1769	2753	3737	787CIP2C_19	3243
786	1770	2754	3738	787CIP2C_20	3243
787	1771	2755	3739	787CIP2C_21	3259
788	1772	2756	3740	787CIP2C_22	3272
789	1773	2757	3741	787CIP2C_23	3278
790	1774	2758	3742	787CIP2C_24	3296
791	1775	2759	3743	787CIP2C_25	3327
792	1776	2760	3744	787CIP2C_26	3334
793	1777	2761	3745	787CIP2C_27	3339
794	1778	2762	3746	787CIP2C_28	3347
795	1779	2763	3747	787CIP2C_29	3387
796	1780	2764	3748	787CIP2C_30	3392
797	1781	2765	3749	787CIP2C_31	3411
798	1782	2766	3750	787CIP2C_32	3427
799	1783	2767	3751	787CIP2C_33	3432
800	1784	2768	3752	787CIP2C_34	3441
801	1785	2769	3753	787CIP2C_35	3479
802	1786	2770	3754	787CIP2C_36	3488
803	1787	2771	3755	787CIP2C_37	3488
804	1788	2772	3756	787CIP2C_38	3553
805	1789	2773	3757	787CIP2C_39	3560
806	1790	2774	3758	787CIP2C_40	3618
807	1791	2775	3759	787CIP2C_41	3642
808	1792	2776	3760	787CIP2C_42	3649
809	1793	2777	3761	787CIP2C_43	3676
810	1794	2778	3762	787CIP2C_44	3747
811	1795	2779	3763	787CIP2C_45	3917
812	1796	2780	3764	787CIP2C_46	4218
813	1797	2781	3765	787CIP2C_47	4219
814	1798	2782	3766	787CIP2C_48	4222
815	1799	2783	3767	787CIP2C_49	4222
816	1800	2784	3768	787CIP2C_50	4229
817	1801	2785	3769	787CIP2C_51	4230
818	1802	2786	3770	787CIP2C_52	4240
819	1803	2787	3771	787CIP2C_53	4241
820	1804	2788	3772	787CIP2C_54	4249
821	1805	2789	3773	787CIP2C_55	4252
822	1806	2790	3774	787CIP2C_56	4267
823	1807	2791	3775	787CIP2C_57	4272
824	1808	2792	3776	787CIP2C_58	4273
825	1809	2793	3777	787CIP2C_59	4275
826	1810	2794	3778	787CIP2C_60	4283
827	1811	2795	3779	787CIP2C_61	4290
828	1812	2796	3780	787CIP2C_62	4292
829	1813	2797	3781	787CIP2C_63	4305
830	1814	2798	3782	787CIP2C_64	4306
831	1815	2799	3783	787CIP2C_65	4308
832	1816	2800	3784	787CIP2C_66	4322
833	1817	2801	3785	787CIP2C_67	4351

834	1818	2802	3786	787CIP2C_68	4356
835	1819	2803	3787	787CIP2C_69	4399
836	1820	2804	3788	787CIP2C_70	4400
837	1821	2805	3789	787CIP2C_71	4520
838	1822	2806	3790	787CIP2C_72	4598
839	1823	2807	3791	787CIP2C_73	4599
840	1824	2808	3792	787CIP2C_74	4600
841	1825	2809	3793	787CIP2C_75	4670
842	1826	2810	3794	787CIP2C_76	4708
843	1827	2811	3795	787CIP2C_77	4734
844	1828	2812	3796	787CIP2C_78	4738
845	1829	2813	3797	787CIP2C_79	4749
846	1830	2814	3798	787CIP2C_80	4752
847	1831	2815	3799	787CIP2C_81	4752
848	1832	2816	3800	787CIP2C_82	4770
849	1833	2817	3801	787CIP2C_83	4784
850	1834	2818	3802	787CIP2C_84	4785
851	1835	2819	3803	787CIP2C_85	4792
852	1836	2820	3804	787CIP2C_86	4803
853	1837	2821	3805	787CIP2C_87	4811
854	1838	2822	3806	787CIP2C_88	4817
855	1839	2823	3807	787CIP2C_89	4818
856	1840	2824	3808	787CIP2C_90	4820
857	1841	2825	3809	787CIP2C_91	4831
858	1842	2826	3810	787CIP2C_92	4841
859	1843	2827	3811	787CIP2C_93	4869
860	1844	2828	3812	787CIP2C_94	4876
861	1845	2829	3813	787CIP2C_95	4902
862	1846	2830	3814	787CIP2C_96	4910
863	1847	2831	3815	787CIP2C_97	4931
864	1848	2832	3816	787CIP2C_98	5303
865	1849	2833	3817	787CIP2C_99	5317
866	1850	2834	3818	787CIP2C_100	5322
867	1851	2835	3819	787CIP2C_101	5330
868	1852	2836	3820	787CIP2C_102	5333
869	1853	2837	3821	787CIP2C_103	5333
870	1854	2838	3822	787CIP2C_104	5356
871	1855	2839	3823	787CIP2C_105	5363
872	1856	2840	3824	787CIP2C_106	5364
873	1857	2841	3825	787CIP2C_107	5379
874	1858	2842	3826	787CIP2C_108	5386
875	1859	2843	3827	787CIP2C_109	5397
876	1860	2844	3828	787CIP2C_110	5401
877	1861	2845	3829	787CIP2C_111	5419
878	1862	2846	3830	787CIP2C_112	5420
879	1863	2847	3831	787CIP2C_113	5452
880	1864	2848	3832	787CIP2C_114	5467
881	1865	2849	3833	787CIP2C_115	5482
882	1866	2850	3834	787CIP2C_116	5483
883	1867	2851	3835	787CIP2C_117	5492
884	1868	2852	3836	787CIP2C_118	5499
885	1869	2853	3837	787CIP2C_119	5525
886	1870	2854	3838	787CIP2C_120	5538
887	1871	2855	3839	787CIP2C_121	5539
888	1872	2856	3840	787CIP2C_122	5558
889	1873	2857	3841	787CIP2C_123	5559
890	1874	2858	3842	787CIP2C_124	5586
891	1875	2859	3843	787CIP2C_125	5619
892	1876	2860	3844	787CIP2C_126	5628
893	1877	2861	3845	787CIP2C_127	5640

894	1878	2862	3846	787CIP2C_128	5640
895	1879	2863	3847	787CIP2C_129	5827
896	1880	2864	3848	787CIP2C_130	6094
897	1881	2865	3849	787CIP2C_131	6195
898	1882	2866	3850	787CIP2C_132	6206
899	1883	2867	3851	787CIP2C_133	6355
900	1884	2868	3852	787CIP2C_134	6362
901	1885	2869	3853	787CIP2C_135	6386
902	1886	2870	3854	787CIP2C_136	6431
903	1887	2871	3855	787CIP2C_137	6457
904	1888	2872	3856	787CIP2C_138	6480
905	1889	2873	3857	787CIP2C_139	6497
906	1890	2874	3858	787CIP2C_140	6532
907	1891	2875	3859	787CIP2C_141	6598
908	1892	2876	3860	787CIP2C_142	6644
909	1893	2877	3861	787CIP2C_143	6644
910	1894	2878	3862	787CIP2C_144	6645
911	1895	2879	3863	787CIP2C_145	6645
912	1896	2880	3864	787CIP2C_146	6761
913	1897	2881	3865	787CIP2C_147	6782
914	1898	2882	3866	787CIP2C_148	6981
915	1899	2883	3867	787CIP2C_149	6981
916	1900	2884	3868	787CIP2C_150	7000
917	1901	2885	3869	787CIP2C_151	7029
918	1902	2886	3870	787CIP2C_152	7885
919	1903	2887	3871	787CIP2C_153	8143
920	1904	2888	3872	787CIP2C_154	8143
921	1905	2889	3873	787CIP2C_155	8234
922	1906	2890	3874	787CIP2C_156	8463
923	1907	2891	3875	787CIP2C_157	8467
924	1908	2892	3876	787CIP2C_158	8540
925	1909	2893	3877	787CIP2C_159	8600
926	1910	2894	3878	787CIP2C_160	9656
927	1911	2895	3879	787CIP2C_161	9669
928	1912	2896	3880	787CIP2C_162	9695
929	1913	2897	3881	787CIP2C_163	9744
930	1914	2898	3882	787CIP2C_164	9849
931	1915	2899	3883	787CIP2D_1	4180
932	1916	2900	3884	787CIP2D_2	4181
933	1917	2901	3885	787CIP2D_3	4314
934	1918	2902	3886	787CIP2D_4	4500
935	1919	2903	3887	787CIP2D_5	5651
936	1920	2904	3888	787CIP2D_6	5691
937	1921	2905	3889	787CIP2D_7	5881
938	1922	2906	3890	787CIP2D_8	5882
939	1923	2907	3891	787CIP2D_9	6209
940	1924	2908	3892	787CIP2D_10	6719
941	1925	2909	3893	787CIP2D_11	8130
942	1926	2910	3894	787CIP2D_12	8863
943	1927	2911	3895	787CIP2D_13	8902
944	1928	2912	3896	787CIP2D_14	9162
945	1929	2913	3897	787CIP2D_15	9197
946	1930	2914	3898	787CIP2D_16	9215
947	1931	2915	3899	787CIP2D_17	9232
948	1932	2916	3900	787CIP2D_18	9262
949	1933	2917	3901	787CIP2D_19	9369
950	1934	2918	3902	787CIP2D_20	9371
951	1935	2919	3903	787CIP2D_21	9516
952	1936	2920	3904	787CIP2D_22	9601
953	1937	2921	3905	787CIP2D_23	9731

954	1938	2922	3906	787CIP2D_24	9733
955	1939	2923	3907	787CIP2D_25	9769
956	1940	2924	3908	787CIP2D_26	9804
957	1941	2925	3909	787CIP2D_27	9816
958	1942	2926	3910	787CIP2D_28	9844
959	1943	2927	3911	787CIP2D_29	9924
960	1944	2928	3912	787CIP2D_30	9936
961	1945	2929	3913	787CIP2D_31	10163
962	1946	2930	3914	787CIP2D_32	10165
963	1947	2931	3915	787CIP2D_33	10165
964	1948	2932	3916	787CIP2D_34	10244
965	1949	2933	3917	787CIP2D_35	10278
966	1950	2934	3918	787CIP2E_1	4251
967	1951	2935	3919	787CIP2E_2	5310
968	1952	2936	3920	787CIP2E_3	5697
969	1953	2937	3921	787CIP2E_4	5731
970	1954	2938	3922	787CIP2E_5	5733
971	1955	2939	3923	787CIP2E_6	5734
972	1956	2940	3924	787CIP2E_7	5740
973	1957	2941	3925	787CIP2E_8	7657
974	1958	2942	3926	787CIP2E_9	9572
975	1959	2943	3927	787CIP2F_1	1363
976	1960	2944	3928	787CIP2F_2	4303
977	1961	2945	3929	787CIP2F_3	5760
978	1962	2946	3930	787CIP2F_4	5766
979	1963	2947	3931	787CIP2F_5	5767
980	1964	2948	3932	787CIP2F_6	5767
981	1965	2949	3933	787CIP2F_7	5770
982	1966	2950	3934	787CIP2F_8	6855
983	1967	2951	3935	787CIP2F_9	10026
984	1968	2952	3936	787CIP2F_10	10227

TABLE 6

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
2953	A	3	324	ISEHRIEASGNYLAQRLTSSFLRGLSSWKSNPLML CGWTILLTLTMVQGEP*GP\KGIPG\FHTNSSYPH WGTVAKPPAGD*DLLPAPGQEGTPLFTR*SLCTY CPID
2954	A	18	467	REELGKDLFDCTLYVLLKYDDFNADKHLALEEF YRAFQVIQLSLPEDQKLSITAATVGQSAVLSCAIQ GTLRPPIIWKRNNIILNNLDLEDINDFGDDGSLYIT KVTTTHVGNYTCYADGYEQVYQTHIFQVNVPPV IRVYPESQARRAG
2955	A	3	23	FYSAFLVADKGIVTSKHNNDTQHIWESDSNEFSV IADPRGNTLGRGTTIT*VSIPPSL
2956	A	1	493	RTKTDVYILNLAVADLLLLFTLPFWAVNAVHGW VLGKIMCKITSALYTLNFVSGMQFLACISIDRYV AVTKVPSQSGVGKPCWIICFCVWMAAILLSIPQL VFYTVNDNARCIPIFPRYLGTSMKALIQMLEICIG FVVPFLIMGVCYFITARTLMKMPNIKIS
2957	A	703	302	EETGVREKRRERMKEKMWQNVLCCTLQTAVIL KLFQNKVLNILKNFFLSPLDTRKNKVFKKWAGG PGAVAHACNPSTLGGRGGRITKSGDRDHPGQHG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ETRSLPACWAQWKSLALPVSRAPGRQGSLVVFPLP
2958	A	575	1054	CTKCKADCDTCFNKNFCTKCKSGFYLHLGKCLDNCPEGLEANNHTMECVSIVHCEVSEWNPWSPCTKKGKTCGFKRGTETRVREIIQHPSAKGNLCPPTNETRKCTVQRKKCQKGERGKKGRERKRKKPNKGESKEAIPDSKSLESSKEIPEQRENKQQQ
2959	A	1	426	LSMLSTISTEHRLSVLWPIWYCCHCPTHLSAVMCVLLWALSLLQSILEWMFCSFLFSDVDSDNWCQILDFLTAVWLIFLI\LVLCGFTLVLLVRIICGSQKMPLTRLYVTILLTGLVFLFCSLPLSIQ*FLLYWIEKDLDDL
2960	A	1194	852	EKRKTSYSQCLNSKQRNVSMRPSIWIHVHLKPPCRLVELLPFSSALQGLSHLSLGTTLP/V*GHLRFRLRNLPQSLRTVILPERNEEQNLQELSHNADKYQMGDCCKEEIDDSIFY
2961	A	274	2250	EKGKVKDAGAEQWISLSLSCKGSWETQFSNHLNSLTPPTSVRRMPLITTVTLLKMVARHHMKLLCSKAFSTQLQQKIFLHSQMGIHHQSVCMKLKPNTSHIISILMGQPMALVQLETLAPLTIIIQKFQTQDHMKFWKNLPLHSHHLTPSVPQTVIPKKTGSPEIKLKITKTIQNGRELFESSLCGDLLNEVQASE\Q*NQSIESRKEKRKKSNKHDSSRSEERKSHKIPKLEPEEQNRPNERVDTVSEKPREEPVLKEGSPSSANTIFCSNNGSVHW\FKFQVGDLVWSKVGTYPWWPCMVSSDPQLEVHTKINTRGAREYHVQFFSNQPERAWVHEKRVREYKGHKQYEELLAEATKQASNHSEKQKIRKPRPQRERAQWDIGIAHAEKALKMTREERIEQYTFIYIDKQPEEALSQAKKSVASKTEVKKTRRPRSVLNTQPEQTNAGEVASSLSSTEIRRHSQRRHTSAEEEEPPPVKIAWKTAAARKSLPASITMHKGSLDLQKCNMSPVVKIEQVFALQNATGDGKFIDQFVYSTKGIGNKTEISVRGQDRLIISTPNQRNEKPTQSVSSPEATSGSTGSVEKKQQRRSIRTRSESEKSTEVVPKKKIKKEQVETVPQATVKTGLQKGSADRGVQGSVRFSDSSVSAAIEETVD
2962	A	2408	836	SASPPPPPPPPPSRFPFSGAPGARDRSGPLGSEPQRNPGARPRTLEATVTPPGSVGAMSSSGLNSEKVAALIQKLNSDPQFVLAQNVGTTHDLLDICLKRATVQRAQHVFQHAVPQEGKPITNQKSSGRCWIFSCLNVMRLPFMKKLNIEEFEFSQSYLFFWDKVERCYFFLSAFVDTAQRKEPEDGRLVQFLLMNPANDGGQWDMLVNIVEKYGVIPKKCFPESYTTEATRRMNDILNHKMREFCIRLRNLVHSGATKGEISATQDVMMEEIFRVVCICLGNPPETFTWEYRDKDKNNKKIGP\ITPLEFNR/EQHVKPLFNMEDKICLVNDPRPQHKYNKLYTV\EYL\SNMVWRGEKLFYNNQPIDFLKKMVAASIKDG\EAVWFGCDVGKHF\NSKLG\LSDMNLYDHELVFGVSLKNMNKAER\LTFGES\LMTHTMTFTAV/SQSRDDSGMVLFTKW\RVGEFQWGEDHGH\KGYLCMTD*VGSLEYVYEVV/VWDRKHVP\EEVLAVLGAGNPFVLPAWDPMGALAE
2963	A	90	543	RHYDSAGKITLKIAKNYLEQRAVGGASPRLAQSVLTCSREPILENSLTSLIEYLHNALEHDMRLRFNNDRMKTTIKETST*LSNSYLVFPLM*SLTYLMKMS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				FERCTARNKMFVNSPFTKVDNYCT\SS\WKKFYL KCYFSLNTIKKEKKMT
2964	A	3	2454	FDTYRGLPSISNGNYSQLQFQAREYSGAPYSQRIS AITTVSVAWKVLSGKIGEGAEGNCKCVISEGAW AVCPTQPCGKAKPDKHLKDLLSKLLNSGYFESIP VPKNAKEKEVPLEEEMLIQSEKKTQLSKTESVKE SESLMEFAQPEIQPQEFLNRRYMTEVDYSNKQGE EQPWEADYARKPNLPKRWDMLTEPDGQEKKQE SFKSWEASGKHQEVSKPAVSLEQRKQDTSKLRS TLPEEQKKQEISKSKPSPSQWKQDTPKSKAGYVQ EEHKKQETPKLWPVQLQKEQDPKKQTPKSWTPS MQSEQNTTKSWTTPMCEEQDSKQPETPKSWENN VESQKHSLTSQSQISPKSWGVATASLIPNDQLLPR KLNTEPKDVP/IACASA*GFLPLQPPFRRI/HVLRK EKLQDLMTQIQGTCNFMQESVLDFDKPSSAIPTS QPPSATPG*PRRHLKEQNLS\VKVIFFQGAVT\VF NVNAPLPPRKEQEIKESPYSPGYNQSFTTASTQTP PQCQLPSIHVEQTVHSQETANYHPDGTIQVSNGS LAFYPAQTNVFPRPTQPFVNSRGSVRGCTRGGRL ITNSYRSPGGYKGFDTYRGLPSISNGNYSQLQFQ AREYSGAPYSQRDNFQQCYKRGGTSGGPRANSR AGWSDSSQVSSPERDNETFNSGDSGQGDSRSMT PVDVPVTNPAATILPVHVYPLPQQMRVAFSAAR TSNLAPGTLDQPIVFDLLLNNLGETFDLQLGRFN CPVNGTYVFIFHMLKLAVNVPLYVNLMKNEEVL VSAYANDGAPDHETASNHAILQLFQGDQIWLRL HRGAIYGSSW
2965	A	3	2454	FDTYRGLPSISNGNYSQLQFQAREYSGAPYSQRIS AITTVSVAWKVLSGKIGEGAEGNCKCVISEGAW AVCPTQPCGKAKPDKHLKDLLSKLLNSGYFESIP VPKNAKEKEVPLEEEMLIQSEKKTQLSKTESVKE SESLMEFAQPEIQPQEFLNRRYMTEVDYSNKQGE EQPWEADYARKPNLPKRWDMLTEPDGQEKKQE SFKSWEASGKHQEVSKPAVSLEQRKQDTSKLRS TLPEEQKKQEISKSKPSPSQWKQDTPKSKAGYVQ EEHKKQETPKLWPVQLQKEQDPKKQTPKSWTPS MQSEQNTTKSWTTPMCEEQDSKQPETPKSWENN VESQKHSLTSQSQISPKSWGVATASLIPNDQLLPR KLNTEPKDVP/IACASA*GFLPLQPPFRRI/HVLRK EKLQDLMTQIQGTCNFMQESVLDFDKPSSAIPTS QPPSATPG*PRRHLKEQNLS\VKVIFFQGAVT\VF NVNAPLPPRKEQEIKESPYSPGYNQSFTTASTQTP PQCQLPSIHVEQTVHSQETANYHPDGTIQVSNGS LAFYPAQTNVFPRPTQPFVNSRGSVRGCTRGGRL ITNSYRSPGGYKGFDTYRGLPSISNGNYSQLQFQ AREYSGAPYSQRDNFQQCYKRGGTSGGPRANSR AGWSDSSQVSSPERDNETFNSGDSGQGDSRSMT PVDVPVTNPAATILPVHVYPLPQQMRVAFSAAR TSNLAPGTLDQPIVFDLLLNNLGETFDLQLGRFN CPVNGTYVFIFHMLKLAVNVPLYVNLMKNEEVL VSAYANDGAPDHETASNHAILQLFQGDQIWLRL HRGAIYGSSW
2966	A	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMGSGI LGLAYVMANTGVFGFSFLLLTVALLASYSVHLL LSMCIQTAYLGP*TNYFMVLPAH*LTCLPLIEFLQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SL*NSL*AVTSYEDLGLFAFGLPGKLVVAGTIIIQ NIGAMSSYLLIIKTELPAAIAEFLTGDYSRYWYLD GQTLLIIICVGIVFPLALLPKIGFLGYTSSLSFFFM MFFALVVIIKKWSIPCPLTLNYVEKGFQISNVTDD CKPKLFHFSKESAYALPTMAFSFLCHTSILPIYCE LQSPSKKRMQNVTNTAIALSFLIYFISALFGYLTF YD/GTTKAQRGEVTCHRIKDKVESELLKG***IP* SHDVVVMT\VKLCILFAVLL\TVPLIHFPARKAVT MMFFSNFPFSWIRHFLITLALNIIIVLLAIYVPDIRN VFGVVGASTSTCLIFIFPGLFYLKLSREDFLSWKK LGVGCFC/LLSFKTSILRNSLSVYIILPASRKSIYFK I
2967	A	3	3222	SGIVVRALWREKKPGGGRRVKRRNPGRQAVGH TEEDPPRVGTPWKEHTGPGPQEGSTMEAAHAKT TEECLAYFGVSETTGLTPDQVKRNLEKYGLNELP AEEGKTLWELVIEQFEDLLVRILLLAACISFVLA WFEEGEETITAFVEPFVILLILIANAIVGVWQERN AENAIEALKEYEPEMGKVYRADRKSVQRIKARD IVPGDIVEVAVGDKVPADIRILAIKSTTLRVDQSIL TGEYVSVIKHTEPVPDPRAVNQDKKNMLFSGTNI AAGKALGIVATTGVGTEIGKIRDQMAATEQDKT PLQQKLDEFGEQLSKVISLICVAVWLINIGHFNDP VHGGSWFRGAIYYFKIAVALAVAAIPEGLPAVIT TCLALGTRRMAKKNAIVRSLPSVETLGCTSVICS DKTGTLTTNQMSVCKMFIIDKVDGDICLLNEFSIT GSTYAPEGEVLKNDKPVRPGQYDGLVELATICA LCNDSSLDFNEAKGVYEKVGEATETALTTLVEK MNVFNTDVRSLSKVERANACNSVIRQLMKKEFT LEFSRDRKSMSVYCSPAKSSRAAVGNKMFVKGA PEGVIDRCNYVRVGTTRVPLTGPVKEKIMAVIKE WGTGRDTLRCLALATRDTPPKREEMVLDDSARF LEYETDLTFVGVVGMLDPPRKEVTGSIQLCRDA GIRVIMITGDNKGTAIAICRRIGIFGENEEVADRA Y\TGREFDDL\PLAEQ\REACRRACCFARVEPSHK SKIVEYLQSYDEITAMTGDGVNDAPALKKAEIGI AMGSGTAVAKTASEMVLADDNFSTIVAAVEEGR AIYNNMKQFIRYLISSNVGEVVCIFLTAALGLPEA LIPVQLLWVNLVTDGLPATALGFNPPDLDIMDRP PRSPKEPLI\SGWLFFRYMAIGGYVGAATVGAAA WWFLYAEDGPHVNYSQLTHFMQCTEDNTHFEGI DCEVFEAPEPMTMALSVLVTIEMCNALNSLSEN QSLLRMPPWVNIWLLGSICLSMSLHFLILYVDPLP MIFKLRALDLTQWLMVLKISLPVIGLDEILKFVA RNYLEG*LFPLLHL*ARVTDPEDERRK
2968	A	3	2414	GARSCSRLGRCTFPLWKGREMEVRKLSISWQFLI VLVLILQILSALDFDPYRVLGVSRTASQADIKKA YKKLAREWHPDKNKDPGAEDKFIQISKAYEILSN EEKRSNYDQYGDAGENQGYQKQQQQREYRFRH FHENFYFDESFFHFPFNSERRDSIDEKYLLHFSHY VNEVAPDSFKKPYLIKITSDWCFSCIHIEPVWKEV IQELEELGVGIGVVHAGYERRLAHHLGAHSTPSI LGIINGKISFFHNAVVRENLRQFVESLLPGNLVEK VTNKNYVRFLSGWQQENKPHVLLFDQTPIVPLL YKLTAFAYKDYLSFGYVYVGLRGTEEMTRRYNI NIYAPTLLVFKEHINRPADVIQARGMKKQIIDDFI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TRNKYLLAARLTSQKLFHELCPVKRSHRQRKYC VVLLTAETTKLSKPFEAFLSFALANTQDTVRFVH VYSNRQQEFADTLLPDSEAFQGKSAVSILERRNT AGRVVYKTLEDPWIGSESDKFILLGYLDQLRKDP ALLSSEAVLPDLTDELAPVFLLRWFYSASDYISD CWDSIFHNNW\REMMPLLSLIFSALFILFGTVIVQ AFSDSNDERESSPPEKEEAQEKTGKTEPSFTKENS SKIPKKGFVEVTELTDVTYTSNLVRLRPGHMNV VLILSNSTKTSLLQKFALEVYTFTGSSCLHFSFLSL DKHREWLEYLLEFAQDAAPIPNQYDKHFMERDY TGYVLALNGHKKYFCLFKPQKTVEEGGKP*GSC SDVDSSLYLGESRGKPSCGLGSRPIKGKLSKLSL WMERLLEGSLQRFYIPSWPELD
2969	A	48	1117	KGLSPDQVLSAFAPLDCEMWLKVFTTFLSFATG ACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTP ASDIQIIWLFERPHTMPKYLLGSVNKSVVPD/YGI P/YTSSP*CHPMASLLINPLQFPDEGNYIVKVNIQG NGTLSASQKIQVTVDDPVTKPVVQIHPPSGAVEY VGNMTLTCHVEGGTRLAYQWLKNGRPVHTSST YSFSPQNNTLHIAPVTKEDIGNYSCLVRNPVSEM ESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTVDL GEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGP RLEVASEKVAQKTMDYVCCAYNNITGRQDETHF TVIITSVGMCDIQGRDPNKT
2970	A	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYADL QFQNSSEMEKIPEIGKFGEKAPPAPSHVWRPAAL FLTLLCLLLLIGLGVLASMFHVTLKIEMKKMNKL QNISEELQRNISLQLMSNMNISNKIRNLSTTLQTI ATKLCRELYSKEQEHKCKPCPRRWIWHKDSCYF LSDDVQTWQESKMACAAQNASLLKINNKNALE FIKSQSRSYDYWLGLSPEEDS/YSWYESG*YNQ\P SAWVIRNAPDLNNMYCGYINRLYVQYYHCTYK QRMICEKMANPVQLGSTYFREA
2971	A	912	2287	VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAPRF LVAFAYWNHYLSCTSPCSCYRPLCRLNFGLNVV ENLALLVLTYVSSSEDF/TWVPG*GRSGEVFPEGT GLPLPHSDLPTSWCGHSLQCGSQSSFPPAIHENAF IVFIASSLGHMLLTCILWRLTKKHTVSQE\DGLSL AGAPRQPRRKSRTSVLRIRVMVRWELSSNGNPG RGVLGLGLGLGNKLRVVGQNLGL*HCVWVVWE TGE*KRWRLQMGIE*GVASRRQ*VRNSVRGLVC HNSSAPPMYMGFFSPTVFGGGVGG*LHVTFILHP PEVEAAGIPLLLGPSLPQRQGREHIVVILAAPACA PFHDR*WEPREIRPSP*ELGLRGEPTLSYPASCRVI RQPIP*DRKSYSWKQRLFIINFISFFSALAVYFRHN MYCEAGVYTIFAILEYTVVLTNMAFHMTAWWD FGNKELLITSQPEEKRF
2972	A	1734	246	GGILSGRDGRTALPRPREPAERTAGLRRDMRPQE LPRLAFPLLLLLLLLLPPPPCPAHSATRFDPTWES LDARQLPAWFDQAKFGIFIHWGVFSVPSFGSEWF WWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPL FTAKFFNANQ\WADIFQASGAKYIVLTSKHHEGF TLWG\SEYSWNWNAIDEGPKRDIVKELEVAIRNR TDLRFGLYYSLFEWFHPLFLEDESSSFHKRQFPVS KTLPELYELVNNYQPEVLWSDGDGGAPDQYWN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				STGFLAWLYNESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGISDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQMGSWLKVNGEAIYETHTWRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWISLEQNGIMVELPQLTIHQMPCKWGWALALTNVI
2973	A	24	1133	SVPRAGGDMETGAAELYDQALLGILQHVGNVQDFLRVLFGFLYRKTDFYRLLRHPSDRMGFPPGAAQALVLQVFKTFDHMARQDDEKRRQELEEKIRRKEEEEAKTVSAAAAEKEPVPVPVQEIEIDSTTELDGHQEVEKVQPPGPVKEMAHGSQEAEAPGAVAGAAEVPR\EPPILPRIQEQFQKNPDSYNGAVRENYTWSQDYTDLEVRVPVPKHVVKGKQVSVALSSSSIRVAMLEENGERVLMEGKLTHKINTESSLWSLEPGKCVLVNLSKVGEYWWNAILEGEEPIDIDKINKERSMATVDEEEQAVLDRLTFDYHQKLQGKPQSHELKVHEMLKKGWDAEGSPFRGQRFDPAMFNISPGAVQF
2974	A	271	1854	MQFGRAHGDCVSGAQLCGCPSMDDYMVLRMIGEGSFGRALLVQHESSNQMFAMKEIRLPKSFSNTQNSRKEAVLLAKMKHPNIVAFKESFEAEGHLYIVMEYCDGGDLMQKIKQQKGKLFPEDMILNWFTQMCLGVNHIHKKRVLHRDIKSKNIFLTQNGKGKLGDFGSARLLSNPMAFACTYVGTPYYVPPEIWENLPYNNKSDIWSLGCILYELCTLKHPFQANSWKNLILKVCQGCISPLPSHYSYELQFLVKQMFKRNPSHRPSATTLLSRGIVARLVQKCLPPEIIMEYGEEVLEEIKNSKHNTPRKKTNPSRIRIALGNEASTVQEEEQDRKGSHTDLESINENLVESALRRVNREEKGNKSVHLRKASSPNLHRRQWEKNVPNTALTALENASILTSSLTAEDDRGGSVIKYSKNTTRKQWLKETPDTLLNILKNADLSLAFQTYTIYRPGS\EGFLKGPLSEETEASDSVDGGHDSVILDPERLEPGLDEEDTDFEEEDDNPDWVSELKKRAGWQGLCDR
2975	A	32	2833	PPGEPGAGRGALSPCGPLSGPPPLPGREAGGTCGQPVNPVFDLSRRNPQEDFELIQRIGSGTYGDVYKARNVNTGELAAIKVIKLEPGEDFAVVQQEIIMMKD\CKHP\DIVAYF\GSYL\RRDKLWI\CMEF\CGSGS\LQDIYHVTGPLSELQIAYVSRETLQGLYYLHSKGKMHRDIKGANILLTDNGHVKLADFGVSAQITATIAKRKSFIGTPYWMAPEVAAVERKGGYNQLCDLWAVGITAIELAELQPPMFDLHPMRALFLMTKSNFQPPKLKDKMKWSNSFHHFVKMALTKNPKKRPTAEKLLQHPFVTQHLTRSLAIELLDKVNNPDHSTYHDFDDDDPEPLVAVPHRIHSTSRNVREEKTRSEITFGQVKFDPPLRKETEPHHELPDSDGFLDSSEEIYYTARSNLDLQLEYGQGHQG\GYFLGANKSLLKSVEEELHQRGHVAHLEDDEGDDDESKHSTLKAKIPPPLPPKPKSIFIPQEMHSTEDENQGTIKRCPMSGSP\AKPSQVPPRPPPPRLPPHKPVALGNGMSSFQLNGERDGSLCQQQNEHRGENLSRKEKKDVPKPISNGLPPTPKVHMGACFSKVFNGCPLKIHCASSWINPDTRDQYLIFGAEEGIYTLNLNELHETSMEQLFPRRCTWLYVMNNCLLSISGKASQLYSHNLPGLFDYA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RQMQKLPVAIPAHKLPDRILPRKFSVSAKIPETK WCQKCCVVRNPYTGHKYLCGALQTSIVLLEWV EPMQKFMLIKHIDFPIPCPLKMFEMLVVPEQEYP LVCVGVSRGRDFNQVVRFETVNPNSTSSWFTES DTPQTNVTHVTQLERDTILVCLDCCIKIVNLQGR LKSSRKLSSELTFDFRIESIVCLQDSVLAFWKHG MQGRSFRSNEVTQEISDSTRIFRLLGSDRVVVLES RPTDNPTANSNLYILAGHENSY
2976	A	32	2833	PPGEPGAGRGALSPCGPLSGPPPLPGREAGGTCG QPVNPVFDLSRRNPQEDFELIQRIGSGTYGDVYK ARNVNTGELAAIKVIKLEPGEDFAVVQQEIIMMK D\CKHP\DIVAYF\GSYL\RRDKLWI\CMEF\CGSGS \LQDIYHVTGPLSELQIAYVSRETLQGLYYLHSKG KMHRDIKGANILLTDNGHVKLADFGVSAQITATI AKRKSFIGTPYWMAPEVAAVERKGGYNQLCDL WAVGITAIELAELQPPMFDLHPMRALFLMTKSNF QPPKLKDKMKWSNSFHHFVKMALTKNPKKRPT AEKLLQHPFVTQHLTRSLAIELLDKVNNPDHSTY HDFDDDDPEPLVAVPHRIHSTSRNVREEKTRSEIT FGQVKFDPPLRKETEPHHELPDSDGFLDSSEEIYY TARSNLDLQLEYGQGHQG\GYFLGANKSLLKSV EEELHQRGHVAHLEDDEGDDDESKHSTLKAKIP PPLPPKPKSIFIPQEMHSTEDENQGTIKRCPMSGSP \AKPSQVPPRPPPPRLPPHKPVALGNGMSSFQLNG ERDGSLCQQQNEHRGENLSRKEKKDVPKPISNG LPPTPKVHMGACFSKVFNGCPLKIHCASSWINPD TRDQYLIFGAEEGIYTLNLNELHETSMEQLFPRR CTWLYVMNNCLLSISGKASQLYSHNLPGLFDYA RQMQKLPVAIPAHKLPDRILPRKFSVSAKIPETK WCQKCCVVRNPYTGHKYLCGALQTSIVLLEWV EPMQKFMLIKHIDFPIPCPLKMFEMLVVPEQEYP LVCVGVSRGRDFNQVVRFETVNPNSTSSWFTES DTPQTNVTHVTQLERDTILVCLDCCIKIVNLQGR LKSSRKLSSELTFDFRIESIVCLQDSVLAFWKHG MQGRSFRSNEVTQEISDSTRIFRLLGSDRVVVLES RPTDNPTANSNLYILAGHENSY
2977	A	174	1543	YSLRKGITFKLAGAMVHIKKGELTQEEKELLEVI GKGTVQEAGTLLSSKNVRVNCLDENGMTPLMH AAYKGKLDMCKLLLRHGADVNCHQHEHGYTA LMFAALSGNKDITWVMLEAGAETDVVNSVGRT AAQMAAFVGQHDCVTIINNFFPRERLDYYTKPQ GLDKEPKLPPKLAGPLHKIITTTNLHPVKIVMLV NENPLLTEEAALNKCYRVMDLICEKCMKQRDM NEVLAMKMHYISCIFQKCINFLKDGENKLDTLIK SLLKG\RASDGFPVYPEKILRESIRK\FPYCEATLL QQLVRSIAPVEIGSDPTAFSVLTQAITGQVGFVDV EFCTTCGEKGASKRCSVCKMVIYCDQTCQKTHW FTHKKICKNLKDIYEKQQLEAAKEKRQEENHGK LDVNSNCVNEEQPEAEVGISQKDSNPEDSGEGK KESLESEAELEGLQDAPAGPQVSEE
2978	A	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNETE DCPGMMLWRYPEPRGLTLVRITPVPFNTTEDPDI STADLGDVLQDPCSLEYWDELQKVFVAFREFNL SESKVCELQLPDINLVNDQKKLVSSDLWRIVLNS SQNGADDQSSASESGSQSTCDPLVTPTALAACTR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VDSCFTPWFVPSLCVSFQFAHLEFHLCHHLDQLG TAAPQYLQPFVSDRNMPSELEYMIVSFREPHMYL RQWNNGSVCQEIQFLAQADCKLLECRNVTMQS VVKPFSIFGQMAVSSDVVEKLLDCTVIVDSVFVN LGQHVVHSLNTAIQAWQQNKCPEVEELVFSHFV ICNDTQETLRFGQVDTDENILLASLHSHQYSWRS HKSPQLLHICIEGWGNWRWSEPFSVDHAGTFIRT IQYRGRTASLIIKVQQLNGVQKQIIICGRQIICSYL SQSIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLESK APEYSIVIQVPSSNSSIIYVWCTVLTLEPNSQVQQ RMIVFSPLFIMRSHLPDPIIIHLEKRSLGLSETQIIP GKGQEKPLQNIEPDLVHHLTFQAREEYDPSDCA VPISTSLIKQIATKVHPGGTVNQILDEFYGPEKSL QPIWPYNKKDSDRNEQLSQWDSPMRVKLSIWKP YVRTLLIELLPWALLINESKWDLWLFEGEKIVLQ VPAGKIIIPPNFQEAFQIGIYWANTNTVHKSVAIK LVHNLTSPKWKDGGNGEVVTLDEEAFVDTEIRL GAFPGHQKLCQFCISSMVQQGIQIIQIEDKTTIINN TPYQIFYKPQLSVCNPHSGKEYFRVPDSATFSICP GGEQPAMKSSSLPCWDLMPDISQSVLDASLLQK' QIMLGFSPAPGADSSQCWSLPAIVRPEFPRQSVA VPLGNFRENGFCTRAIVLTYQEHLGVTYLTLSED PSPRVIIHNRCPVKMLIKENIKDIPKFEVYCKKIPS ECSIHHELYHQISSYPDCKTKDLLPSLLLRVEPLD EVTTEWSDAIDINSQGTQVVFLTGFGYVYVDVV HQCGTVFITVAPEGKAGPILTNTNRAPEKIVTF/K MFITQLSLAVFDDLTHHKASAELLRLTLDNIFLC VAPGAGPLPGEEPVAALFELYCVEICCGDLQLDN QLYNKSNFHFAVLVCQGEKAEPIQCSKMQSLLIS NKELEEYKEKCFIKLCITLNEGKSILCDINEFSFEL KPARLYVEDTFVYYIKTLFDTYLPNSRLAGHSTH LSGGKQVLPMQVTQHARALVNPVKLRKLVIQPV NLLVSIHASLKLYIASDHTPLSFSVFERGPIFTTAR QLVHALAMHYAAGALFRAGWVVGSLDILGSPA SLVRSIGNGVADFFRLPYEGLTRGPGAFVSGVSR GTTSFVKHISKGTLTSITNLATSLARNMDRLSLDE EHYNRQEEWRRQLPESLGEGLRQGLSRLGISLLG AIAGIVDQPMQNFQKTSEAQASAGHKAKGVISG VGKGIMGVFTKPIGGAAELVSQTGYGILHGAGLS QLPKQRHQPSD\VHADQAPNSHVKYVWKMLQS LGRPEVHMALDVVLVRGSGQEHEGCLLLTSEVL FVVSVSEDTQQQAFPVTEIDCAQDSKQNNLLTV QLKQPRVACDVEVDGVRERLSEQQYNRLVDYIT KTSCHLAPSCSSMQIPCPVVAAEPPPSTVKTYHY LVDPHFAQVFLSKFTMVKNKALRKGFP
2979	A	255	2673	AWLFPASVLCPRCLTGSAVGSAEWKSLVVLFPFS SRPTLGHLDSKPSSKSNMIRGRNSATSADEQPHIG NYRLLKTIGKGNFAKVKLARHILTGKEVAVKIID KTQLNSSSLQKLFREVRIMKVLNHPNIVKLFEVIE TEKTLYLVMEYASGGEVFDYLVAHGRMKEKEA RAKFRQIVSAVQYCHQKFIVHRDLKAENLLLDA DMNIKIADFGFSNEFTFGNKLDTFCGSPPYAAPEL FQGKKYDGPEVDVWSLGVILYTLVSGSLPFDGQ NLKELRERVLRGKYRIPFYMSTDCENLLKKFLIL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NPSKRGTLEQIMKDRWMNVGHE\DDELKPYGEP LP\DYKDPRRTELMVSMGYTREEIQDSLVGQRYN EVMATYLLLGYKSSELEGDTITLKPRPSADLTNS SAPSPSHKVQRSVSANPKQRRFSDQAGPAIPTSNS YSKKTQSNNAENKRPEEDRESGRKASSTAKVPA SPLPGLERKKTTPTPSTNSVLSTSTNRSRNSPLL\E RASL\GQGFHPEWAKTALTMPGSRASTASASAA VSAARPRQHQKSMSASVHPNKASGLPPTESNCE VPRPRQVCWGSCTAPQRVPVASPSAHNISSSGGA PDRTNFPRGVSSRSTFHAGQLRQVR\DQQNLPYG VTPASPSGHSQGRRGASGSIFSKFTSKFVRRNLNE PESKDR\VETLRPHVV\NSGGNDKEKEEFREAKPR SLRFTWSMKTTSSMEPNEMMREIRKVLDANSCQ SELHEKYMLLCMHGTPGHEDFVQWEMEVCKLP RLSLNGVRFKRISGTSMAFKNIASKIANELKL
2980	A	120	3433	NCLLLQAKGFHGEIEDLQQWLTDTERHLLASKP LGGLPETAKEQLNVHMEVCAAFEAKEETYKSLM QKGQQMLARCPKSAETNIDQDINNLKEKWESVE TKLNER\KT\KLEEALNLA\MEFHNSL\QDFINWLT QAEQTLNVASRPSLILDTVLFQIDEHKVFANEVN SHREQIIELDKTGTHLKYFSQKQDVVLIKNLLISV QSRWEKVVQRLVERGRSLDDARKRAKQFHEAW SKLMEWLEESEKSLDSELEIANDPDKIKTQLAQH KEFQKSLGAKHSVYDTTNRTGRSLKEKTSLADD NLKLDDMLSELRDKWDTICGKSVERQNKLEEA\ LLFSGQFTDALQALIDWLYRVEPQLAEDQPVHG DIDLVMNLIDNHKAFQKELGKRTSSVQALKRSA RELIEGSRDDSSWVKVQMQELSTRWETVCALSIS KQTRLEAALRQAEEFHSVVHALLEWLAEAEQTL RFHGVLPDDEDALRTLIDQHKEFMKKLEEKRAE LNKATTMGDTVLAICHPDSITTIKHWITIIRARFEE VLAWAKQHQQRLASALAGLIAKQELLEALLAW LQWAETTLTDKDKEVIPQEIEEVKALIAEHQTFM EEMTRKQPDVDKVTKTYKRRAADPSSLQSHIPV LDKGRAGRKRFPASSLYPSGSQTQIETKNPRVNL LVSKWQQVWLLALERRRKLNDALDRLEELREF ANFDFDIWRKKYMRWMNHKKSRVMDFFRRIDK DQDGKITRQEFIDGILSSKFPTSRLEMSAVADIFD RDGDGYIDYYEFVAALHPNKDAYKPITDADKIE DEVTRQVAKCKCAKRFQVEQIGDNKYRFFLGNQ FGDSQQLRLVRILRSTVMVRVGGGWMALDEFL VKNDPCRAKGRTNMELREKFILADGASQGMAA FRPRGRRSRPSSRGASPNRSTSVSSQAAQAASPQ VPATTTPKILHPLTRNYGKPWLTNSKMSTPCKAA ECSDFPVPSAEGTPIQGSKLRLPGYLSGKGFHSGE DSGLITTAAARVRTQFADSKKTPSRPGSRAGSKA GSRASSRRGSDASDFDISEIQSVCSDVETVPQTHR PTPRAGSRPSTAKPSKIPTPQRKSPASKLDKSSKR
2981	A	120	3433	NCLLLQAKGFHGEIEDLQQWLTDTERHLLASKP LGGLPETAKEQLNVHMEVCAAFEAKEETYKSLM QKGQQMLARCPKSAETNIDQDINNLKEKWESVE TKLNER\KT\KLEEALNLA\MEFHNSL\QDFINWLT QAEQTLNVASRPSLILDTVLFQIDEHKVFANEVN SHREQIIELDKTGTHLKYFSQKQDVVLIKNLLISV QSRWEKVVQRLVERGRSLDDARKRAKQFHEAW

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SKLMEWLEESEKSLDSELEIANDPDKIKTQLAQHKEFQKSLGAKHSVYDTTNRTGRSLKEKTSLADDNLKLDDMLSELRDKWDTICGKSVERQNKLEEA\LLFSGQFTDALQALIDWLYRVEPQLAEDQPVHGDIDLVMNLIDNHKAFQKELGKRTSSVQALKRSARELIEGSRDDSSWVKVQMQELSTRWETVCALSISKQTRLEAALRQAEEFHSVVHALLEWLAEAEQTLRFHGVLPDDEDALRTLIDQHKEFMKKLEEKRAELNKATTMGDTVLAICHPDSITTIKHWITIIRARFEEVLAWAKQHQQRLASALAGLIAKQELLEALLAWLQWAETTLTDKDKEVIPQEIEEVKALIAEHQTFMEEMTRKQPDVDKVTKTYKRRAADPSSLQSHIPVLDKGRAGRKRFPASSLYPSGSQTQIETKNPRVNLLVSKWQQVWLLALERRRKLNDALDRLEELREFANFDFDIWRKKYMRWMNHKKSRVMDFFRRIDKDQDGKITRQEFIDGILSSKFPTSRLEMSAVADIFDRDGDGYIDYYEFVAALHPNKDAYKPITDADKIEDEVTRQVAKCKCAKRFQVEQIGDNKYRFFLGNQFGDSQQLRLVRILRSTVMVRVGGGWMALDEFLVKNDPCRAKGRTNMELREKFILADGASQGMAAFRPRGRRSRPSSRGASPNRSTSVSSQAAQAASPQVPATTTPKILHPLTRNYGKPWLTNSKMSTPCKAAECSDFPVPSAEGTPIQGSKLRLPGYLSGKGFHSGEDSGLITTAAARVRTQFADSKKTPSRPGSRAGSKAGSRASSRRGSDASDFDISEIQSVCSDVETVPQTHRPTPRAGSRPSTAKPSKIPTPQRKSPASKLDKSSKR
2982	A	1	2065	MAAGGAEGGSGPGAAMGDCAEIKSQFRTREGFYKLLPGDGAARRSGPASAQTPVPPQPPQPPPGPASASGPGAAGPASSPPPAGPGPGPALPAVRLSLVRLGEPDSAGAGEPPATPAGLGSGGDRVCFNLGRELYFYPGCCRRGSQRWHTPLTPFLPPLKSIDLNKPIDKRIYKGTQPTCHDFNQFTAATETISLLVGFSAGQVQYLDLIKKDTSKLFNEERLIDKTKVTYLKWLPESESLFLASHASGHLYLYNVSHPCASAPPQYSLLKQ\AWGFSFYAAKSKAPRNPLAKWAVGEGPLNEFAFSPDGRHLACVSQDGCLRVFHFDSMLLRGLMKSYFGGLLCVCWSPDGRYVVTGGEDDLVTVWSFTEGRVVARGHGHKSWVNAVAFDPYTTRAEEAATAAGADGERSGEEEEEEPEAAGTGSAGGAPLSPLPKAGSITYRFGSAGQDTQFCLWDLTEDVLYPHPPLARTRTLPGTPGTTPPAASSSRGGEPGPGPLPRSLSRSNSLPHPAGGGKAGGPGVAAEPGTPFSIGRFATLTLQERRDRGAEKEHKRYHSLGNISRGGSGGSGSGGEKPSGPVPRSRLDPAKVLGTALCPRIHEVPLLEPLVCKKIAQERLTVLLFLEDCIITACQEGLICTWARPGKAFTDEETEAQTGEGSWPRSPSKSVVEGISSQPGNSPSGTVV
2983	A	3855	220	RRFRLSAHRAQPCCRCRGLEMPRGVFQQLSNLVLQELNANLSNLTSAFEKATAEKIKCQQEADATNRVILLANRLVGGLASENIRWAESVENFRSQGVTLCGDVLLISAFVSYVGYFTKKYRNELMEKFWIPYIHNLKVPIPITNGLDPLSLLTDDADVATWNNQGLPSDRMSTENATILGNTERWPLIVDAQLQGIKWIKNKYRSELKAIRLGQKSYLDVIEQATSEGDTLLIENIGETVDPALDPLLGRNTIKKGKYIKIGDKEVGVPP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QVPPDPTHQVLQPTLQARDAGSVH\LINFLVTRD GLEDQLLAAVVAKERPDLEQLKANLTKSQNEFK IVLKELEDSLLARLSAASGNFLGDTALVENLETT KHTASEIEEKVVEAKITEVKINEARENYRPAAER ASLLYFILNDLNKINPVYQFSLKAFNVVFEKAIQR TTPANEVKQRVINLTDEITYSVYMYTARGLFERD KLIFLAQVTFQVLSMKKELNPVELDFLLRFPFKA GVVSPVDFLQHQGWGGIKALSEMDEFKNLDSDI EGSAKRWKKLVESEAPEKEIFPKEWKNKTALQK LCMVRCLRPDRMTYAIKNFVEEKMGSKFVEGRS VEFSKSYEESSPSTSIFFILSPGVDPLKDVEALGKK LGFTIDNGKLHNVSLGQGQEVVAENALDVAAEK GHWVILQNIHLVARWLGTLDKKLERYSTGRHED YRVFIRAEPAPSPETHIIPQGILENAIKITNEPPTGM YANLYKALDLFTQDTLEMCTKEMEFKCMLFAL CYFHAVVAERRKFGAQGWNRSYPFNNGDLTISI NVLYNYLEANPKVPWDDLRYLFGEIMYGGHITD DWDRRLCRTYLAEYIRTEMLEGDVLLAPGFQIPP NLDYKGYHEYIDENLPPESPYLYGLHPNAEIGFL TVTSEKLFRTVLEMQPKETDSGAGTGVSREEKV KAVLDDILEKIPETFNMAEIMAKAAEKTPYVVV AFQECERMNILTNEMRRSLKELNLGLKGELTITT DVEDLSTALFYDTVPDTWVARAYPSMMGLAAW YANLLLRIRELEAWTTDFALPTTVWLAGFFNPQS FLTAIMQSMARKNEWPLDKMCLSVEVTKKNRE DMTAPPREGSYVYGLFMEGARWDTQTGVIAEA RLKELTPAMPVIFIKAIPVARMETKNIYECPVYKT RIRGPTYVWTFNLKTKEKAAKWILAAVALLLQV
2984	A	2	1464	FVLFPGIAMETPGASASSLLLPAASRPPRKREAGE AGAATSKQRVLDEEEYIEGLQTVIQRDFFPDVEK LQAQKEYLEAEENGDLERMRQIAIKFGSALGKM SREPPPPYVTPATFETPEVHAGTGVVGNKPRPRG RGLEDGEAGEEEEKEPLPSLDVFLSRYTSEDNAS FQEIMEVAKERSRARHAWLYQAEEEFEKRQKDN LELPSAEHQAIESSQASVETWKYKAKNSLMYYP EGVPDEEQLFKKPRQVVHKNTRFLRDPFSQALSR CQLQQAAALNAQHKQGKVGPDGKELIPQESPRV GGFGFVATPSPAPGVNESPMMTWGEVENTPLRV EGSETPYVDRTPGPAFKILEPGRRERLGLKMANE AAAKNRAKKQEALRRVTENLASLTPKGLSPAMS PALQRLVSRTASKYTDRALRASYTPSPARSTHLK NPGPVGCRPPQSTPGA/PGSATRTPL\TQDPA\SIT DNLLQLPARRKASDFF
2985	A	1890	178	ASTQEAGLLSPPGVGAQRCWNFVACLPVRACAD MASNDYTQQATQSYGAYPTQPGQGYSQQSSQP YGQQSYSGYSQSTDTSGYGQSSYSSYGQSQNSY GTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGY GQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQQPS YGGQQQSYGQQQSYNPPRGYGQQNQYNSSSGG GGGGGGGGSYGQDQSSMSGSGGGGGGGGGGS GGGGGYGNQDQTGAAGSRGYRQ\QDRGGRCRG GSGGGGS\GGAAGYNRSSGGYEPRGRGGGRGGR GGMGGSDRGGFNKFGGPRDQGSRHDSEQDNSD NNTIFVQGLGENVTIESVADYFKQIGIIKTNKKTG QPMINLYTDRETGKLKGEATVSFDDPPSAKAAID

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				WFDGKEFSGNPIKVSFATRRADFNRGGGNGRGG RGRGGPMGRGGYGGGGSGGGGRGGFPSGGGGG GGQQRAGDWKCPNPTCENMNFSWRNECNQCK APKPDGPGGGPGGSHMGGNYGDDRRGGRGGYD RGGYRGRGGDRGGFRGGRGGGDRGGFGPGKM DSRGEHRQDRRERPY
2986	A	1890	178	ASTQEAGLLSPPGVGAQRCWNFVACLPVRACAD MASNDYTQQATQSYGAYPTQPGQGYSQQSSQP YGQQSYSGYSQSTDTSGYGQSSYSSYGQSQNSY GTQSTPQGYGSTGGYGSSQSSQSSYGQQSSYPGY GQQPAPSSTSGSYGSSSQSSSYGQPQSGSYSQQPS YGGQQQSYGQQQSYNPPRGYGQQNQYNSSSGG GGGGGGGGSYGQDQSSMSGSGGGGGGGGGGS GGGGGYGNQDQTGAAGSRGYRQ\QDRGGRCRG GSGGGGS\GGAAGYNRSSGGYEPRGRGGGRGGR GGMGGSDRGGFNKFGGPRDQGSRHDSEQDNSD NNTIFVQGLGENVTIESVADYFKQIGIIKTNKKTG QPMINLYTDRETGKLKGEATVSFDDPPSAKAAID WFDGKEFSGNPIKVSFATRRADFNRGGGNGRGG RGRGGPMGRGGYGGGGSGGGGRGGFPSGGGGG GGQQRAGDWKCPNPTCENMNFSWRNECNQCK APKPDGPGGGPGGSHMGGNYGDDRRGGRGGYD RGGYRGRGGDRGGFRGGRGGGDRGGFGPGKM DSRGEHRQDRRERPY
2987	A	1376	898	GGAKAGGAPHPFTLPFRHVGGLSAAPEEVEGML WAGARQHGRNWRKRETSPGTQGPLPPVPR/VPP GPDG\PHAIAPTLSWAIPRQQCSPQPGRLNALPPD RCSGPHFGDRAPESCFPGACSVSGACAFKGTRPA CPPQEPSLRSSRNRLREGQTFGRMEI
2988	A	1	1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAIDP LRVAPLPLYAAIFLVGVPGNAMVAWVAGKVAR RRVGATWLLHLAVADLLCCLSLPILAVPIARGGH WPYGAVGCRALPSIILLTMYASVLLLAALSADLC FLALGPAW\CLRFS/GACGVQVACGAAWTLALL LTVPSAIYRRLHQEHFPARLQCVVDYGGSSSTEN AVTAIRFLFGFLGPLVAVASCHSALLCWAARRC RPLGTAIVVGFFVCWAPYHLLGLVLTVAAPNSA LLARALRAEPLIVGLALAHSCLNPMLFLYFGRAQ LRRSLPAACHWALRESQGQDESVDSKKSTSHDL VSEMEV
2989	A	27	4074	KSQLFCFWVGKAGDILSGDQDKEQKDPYFVETP YGYQLDLDFLKYVDDIQKGNTIKRLNIQKRRKPS VPCPEPRTTSGQQGIWTSTESLSSSNSDDNKQCP NFLIARSQVTSTPISKPPPPLETSLPFLTIPENRQLP PPSPQLPKHNLHVTKTLMETRRRLEQERATMQM TPGEFRRPRLASFGGMGTTSSLPSFVGSGNHNPA KHQLQNGYQGNGDYGSYAPAAPTTSSMGSSIRH SPLSSGISTPVTNVSPMHLQHIREQMAIALKRLKE LEEQVRTIPVLQVKISVLQEEKRQLVSQLKNQRA ASQINVCGVRKRSYSAGNASQLEQLSRARRSGG ELYIDYEEEEMETVEQSTQRIKEFRQL\TADMQA LEQKIQDSSCEASSELRENGECRSVAVGAEENMN DIVVYHRGSRSCKDAAVGTLVEMRNCGVSVTEA MLGVMTEADKEIELQQQTIESLKEKIYRLEVQLR ETTHDREMTKLKQELQAAGSRKKVDKATMAQP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LVFSKVVEAVVQTRDQMVGSHMDLVDTCVGTS VETNSVGISCQPECKNKVVGPELPMNWWIVKER VEMHDRCAGRSVEMCDKSVSVEVSVCETGSNTE ESVNDLTLLKTNLNLKEVRSIGCGDCSVDVTVCS PKECASRGVNTEAVSQVEAAVMAVPRTADQDT STDLEQVHQFTNTETATLIESCTNTCLSTLDKQTS TQTVETRTVAVGEGRVKDINSSTKTRSIGVGTLL SGHSGFDRPSAVKTKESGVGQININDNYLVGLK MRTIACGPPQLTVGLTASRRSVGVGDDPVGESLE NPQPQAPLGMMTGLDHYIERIQKLLAEQQTLLA ENYSELAEAFGEPHSQMGSLNSQLISTLSSINSVM KSASTEELRNPDFQKTSLGKITGSYLGYTCKCGG LQSGSPLSSQTSQPEQEVGTSEGKPISSLDAFPTQ EGTLSPVNLTDDQIAAGLYACTNNESTLKSIMKK KDGNKDSNGAKKNLQFVGINGGYETTSSDDSSS DESSSSESDDECDVIEYPLEEEEEEEDEDTRGMAE GHHAVNIEGLKSARVEDEMQVQECEPEKVEIRE RYELSEKMLSACNLLKNTINDPKALTSKDMRFC LNTLQHEWFRVSSQKSAIPAMVGDYIAAFEAISP DVLRYVINLADGNGNTALHYSVSHSNFEIVKLLL DADVCNVDHQNKAGYTPIMLAALAAVEAEKDM RIVEELFGCGDVNAKASQAGQTALMLAVSHGRI DMVKGLLACGADVNIQDDEGSTALMCASEHGH VEIVKLLLAQPGCNGHLEDNDGSTALSIALEAGH KDIAVLLYAHVNFAKAQSPGTPRLGRKTSPGPTH RGSFD
2990	A	69	1687	ERLRPGQRAIRGPVPAAGACASLPPRAGPAQGRH AALGGAEPGSHLHCGVRLQRREEPGGQQRLLPQ RGGSAQTGHQHPGPYECQCPGPQPGGTTPALLSL ILEETRGPPASANPDKDHSTQPGTMGRKKIQISRI LDQRNRQVTFTKRKFGLMKKAYELSVLCDCEIA LIIFNSATRLFQYASTDMDRVLLKYTEYSEPHESR TNTDILETLKRRGIGLDGPELEPDEGPEEPGEKFR RLAGEGGDPALPRPRLYPAAPAMPSPDVVYGAL PPPG\CDPSGLGEALPAQSRPSPFRPAAPKAGPPG LGHPLFSPSHLTSKTPPPLYLPTEGRRSDLPGGLA GPRGGLNTSRSLYSGLQNPCSTATPGPPLGSFPFL PGGPPVGAEAWARRVPQPAAPPRRPPQSSIKSER LFLRPPGAPATFLRPSPIPCSSPGPWQSLCGLGPP\ CAGCPWPTAGPGRRSPGGTSPERSPGTARARGDP \TSLQAFSEKTHTVTAPLRGGGLEVGGWTQSSAG GLLSFFLFVCISTNKNARGVRGPEKK
2991	A	3	1159	IPQPLHCASPKEEMSLRCGDAARTLGPRVFGRYF CSPVRPLSSLPDKKKELLQNGPDLQDFVSGDLAD RSTWDEYKGNLKRQKGERLRLPPWLKTEIPMGK NYNKLKNTLRNLNLHTVCEEARCPNIGECWGGG EYATATATIMLMGDTCTRGCRFCSVKTARNPPP LDASEPYNTAKAIAEWGLDYVVLTSVDRDDMP DGGAEHIAKTVSYLKERNPKILVECLTPDFRGDL KAIEKVALSGLDVYAHNVETVPELQSKVRDPRA NFDQSLRVLKHAKKVQPDVISKTSIMLGLGENDE QVYATMKALREADVDCLTLGQYMQPTRRHLKV EEYITPEKFKYWEKVGNELGFHYTASGP\LVRSS YKAGEFFLKNLVAKRKTKDL
2992	A	3	1636	PVPGVPTSPPSCCPQDMQGPWVLLLLGLRLQLSL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GVIPAEEENPAFWNRQAAEALDAAKKLQPIQKVAKNLILFLGDGLGVPTVTATRILKGQKNGKLGPETPLAMDRFPYLALSKTYNVDRQVPDSAATATAYLCGVKANFQTIGLSAAARFNQCNTTRGNEVISVMNRAKQAGKSVGVVTTTRVQHASPAGTYAHTVNRNWYSDADMPASARQEGCQDIATQLISNMDIDVILGGGRKYMFPMGTPDPEYPADASQNGIRLDGKNLVQEWLAKHQGAWYVWNRTELMQASLDQSVTHLMGLFEPGDTKYEIHRDPTLDPSLMEMTEAALRLLSRNPRGFYLFVEGGRIDHGHHEGVAYQALTEAVMFDDAIERAGQLTSEEDTLTLVTADHSHVFSFGGYTLRGSSIFGLAPSKAQDSKAYTSILYGNGPGYVFNSGVRPDVNESESGSPDYHQQAG\VPLSSETHGGEDVAVFARGPQAHLVHGVQEQSFVAHVMAFAACLEPYTACDLAPPACTTDAAHPVAASLPLLAGTLLLLGASAAP
2993	A	3	685	DAWARLLKMNRLFGKAKPKAPPPSLTDCIGTVDSRAESIDKKISRLDAELVKYKDQIKKMREGPAKNMVKQKALRVLKQKRMYEQQRDNLA\NSHSTW\TS\HYTIQSLKDTKTTVDAMKLGVKEMKKAYKQVKIDQIEDLQDQLEDMMEDANEIQEALSRSYGTPELDEDDLEAELDALGDELLADEDSSYLDEAASAPAIPEGVPTDTKNKDGVLVDEFGLPQIPAS
2994	A	1710	161	RRCELTPFIIKTLILPKSWGAFPEDVVMQHVSSSQSSQRHVQWPGACPGAGEEQPACSQPSLPLTLPSPSHQLQQLMVRGGPAGGQNMNVDLQGVGPGLQGSPQVTLAPLPLPSPTSPGFQFSAQPRRFEHGSPSYIQVTSPLSQQVQTQSPTQPSPGPGQALQNVRAGAPGPGLGLCSSSPTGDFVDASVLVRQISLSPSSGGHFVFQDGSGLTQIAQGAQVQLQHPGTPITVRERRPSQPHTQSGGTIHHLGPQSPAAAGGAGLQPLASPSHITTANLPPQISSIIQGQLVQQQQVLQGPPLPRPLGFERTPGVLLPGAGGAAGFGMTSPPPPTSPSRTAVPPGLSSLPLTSVGNTGMKKVPKKLEEIPPASPEMAQMRKQCLDYHHQEMQALKEVFKEYLIELFFLQHFQGNMMDFLAFKERLYGPLQAYLRQNDLDIEEEEEE\HFEVINDEVKVVARKHGQPGTPVAIAT\QLPPRTSAAFPAQQQPLQVLSDGSTVQLPRLSSLGFEDSMC
2995	A	3	924	SAPSGIDASTHAFARCKHPINVRRDPSIPIYGLRQSILLNTRLQDCYVDSPALTNIWMARTCAKQNINAPAPATTSSWEVVRNPLIASSFSLVKLVLRRQLKNKCCPPPCKFGEGKLSKRLKHKDDSVMKATQQARKRNFISSKSKQPAGHRRPAGGIRESKESSKEKKLTVRQDLEDRYAEHVAAT\QALPQDSGTAAWKG\RVLLPETQKRQQLSEDTLTIHGLPTEGYQALYHAVVEPMLWNPSGTPKRYSLELGKAIKQKLWEALCSQGAISEGAQRDRFPGRKQPGVHEEPVLKKWPKLKSKK
2996	A	3	1713	GKFGIKPSQRRISGKSTFHSEMEGEDTRDDSLYSILEELWQDAEQIKRCQEKHNKLLSRTTFLNKKILNTEWDYEYKDFGKFVHPSPNLILSQKRPHKRDSFGKSFKHNLDLHIHNKSNAAKNLDKTIGHGQVFTQNSSYSHHENTHTGVKFCERNQCGKVLSLKHSLSQNVKFPIGEKANTCTEFGKIFTQRSHFFAPQKIHT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VEKPHELSKCVNVFTQKPLLSIYLRVHRDEKLYI\CTKM/CGKGLHPRNSELIMHEKTHTREKPYKCNE\CGKSFFQVSSLLRHQTTHTGEKLFECSECGKGFSLNSALNIHQKIHTGERHHKCSECGKAFTQKSTLRMHQRIHTGERSYICTQCGQAFIQKAHLIAHQRIHTGEKPYECSDCGKSFPSKSQLQMHKRIHTGEKPYICTECGKAFTNRSNLNTHQKSHTGEKSYICAECGKAFTDRSNFNKHQTIHTGEKPYVCADCGRAFIQKSELITHQRIHTTEKPYKCPDCEKSFSKKPHLKVHQRIHTGEKPYICAECGKAFTDRSNFNKHQTIHTGDKPYKCSDCGKGFTQKSVLSMHRNIHT
2997	A	3	1763	AASTRTMGSRHFEGIYDHVGHFGRFQRVLYFICAFQNISCGIHYLASVFMGVTPHHVCRPPGNVSQVVFHNHSNWSLEDTGALLSSGQKDYVTVQLQNGEIWELSRCSRNKRENTSSLGYEYTGSKKEFPCVDGYIYDQNTWKSTAVTQWNLVCDRKWLAMLIQPLFMFGGPTGIG/VTFGYF\SDRLGRRVVLWATSSSMFLFGIAAAFAVDYYTFMAARFFLAMVASGYLVVGFVYVMEFIGMKSRTWASVHLHSFFAVGTLLVALTGYLVRTWWLYQMILSTVTVPFILCCWVLPETPFWLLSEGRYEEAQK\IVDIMAKWNRASSCKLSELLSLDLQGPVSNSPTEVQKHNLSYLFYNWSITKRTLTVWLIWFTGSLGFYSFSLNSVNLGGNEYLNLFLLGVVEIPAYTFVCIAMDKVGRRTVLAYSLFC\SALACGVVMVIPQKHYILGVVTAM\VGKILPIGAAFG\LIYLYTAELYPTIVRSLAVGSGSMVCRLASILAPFSVDLSSIWIFIPQLFVGTMALLSGVLTLKLPETLGKRLATTWEEAAKLESENESKSSKLLLTTNNSGLEKTEAITPRDSGLGE
2998	A	3	1441	QRPASQLLAPFAAEALPGAPRAAMAQHFSLAACDVVGFDLDHTLCRYNLPESAPLIYNSFAQFLVKEKGYDKELLNVTPEDWDFCCKGLALDLEDGNFLKLANNGTVLRASHGTKMMTPEVLAEAYGKKEWKHFLSDTGMACRSGKYYFYDNYFDLPGALLCARVVDYLTKLNNGQKTFDFWKDIVAAIQHNYKMSAFKENCGIYFPEIKRDPGRYLHSRPESVKKWLRQLKNAGKILLLITSSHSDYCRLLCA\YILGNDFTDLFDIVITNALKPGFFSHLPSQRPFRTLENDEEQEALPSLDKPGWYSQGNAVHLYELLKKMTGKPEPKVVYFGDSMHSDIFPARHYSNWETVLILEELRGDEGTRSQRPEESEPLEKKGKYEGPKAKPLNTSSKKWGSFF\IDSVLGLENTEDSLVYTWSCKRISTYSTIAIPSIEAIAELPLDYKFTRFSSSNSKTAGYYPNPPLVLSSDETLISK
2999	A	320	2417	LRRRKMTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSEEALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSDKASSLLCFQHQEESLAQGPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTGAHNASVDMCELKRDLQLLSQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMGSFEEDRINATVWKLQPTAGLQDLHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLLVDFSSQALFQDKNSSQVLGEKVLGIVVQNTKVANLTEPVVLTFQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGC

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ETVRRETQTSCFCNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPLPCRRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLEGYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIYPSMCWIRDSLVSYITNLGLFSLVFLFNMAMLATMVVQILRLRPHTQKWSHVLTLLCLSLVLG\LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLPISSGSTSSSRI
3000	A	66	1003	SRGQLDAGQSSEQHGGNRQPEQSRSRSSSSSSSPRRSRSAAEPAMALSMPLNGLKEEDKEPLIELFVKAGSDGESIGNCPFSQRLFMILWLKGVVFSVTTVDLKRKPADLQNLAPGTHPPFITFNSEVKTDVNKIEEFLEEVLCPPKYLKLSPKHPESNTAGMDIFAKFSAYIKNSRPEANEALERGLLKTLQKLDEYLNSPLPDEIDENSMEDIKFSTRKFLDGNEMTLADCNLLPKLHIVKVVAKKYRNFDIPKEMTGIWRYLTNAYSRDEFTNTCPSDKEVEI\AYSDVAKRLHQVKSRLLKEVSFMSSP
3001	A	779	2006	LALTFRSALSTLPGSPMTSSGSPDLQLAWGPSLLPHPPSVWSPALPSCFAGPCPLLPLSDTQGWWGPNWLAPPSAALCRPDAAVWPDLPSSNILLVTPPPAK*SAVAV*PCPRGAHSLERAARQYTISGSSTSQSGKCSKRDTKCCAVTTSWGCFWQKHWKGDEDSGWAFQEGSHLGEGHL
3002	A	909	2799	VEEAWTVWLHWGVRECLLEEETNQKEEAASSNWTKARGPFWQEDWVWDMRLKMTTRNFPEREVPCDVEVERFTREVPCLSSLGDGWDCENQEGHLRQSALTLEKPGTQEAICEYPGFGEHLIASSDLPPSQRVLATNGFHAPDSNVSGLDCDPALPSYPKSYADKRTGDSDACGKGFNHSMEVIHGRNPVREKPYKYPESVKSFNHFTSLGHQKIMKRGKKSYEGKNFENIFTLSSSLNENQRNLPGEKQYRCTECGKCFKRNSSLVLHHRTHTGEKPYTCNECGKSFSKNYNLIVHQRIHTGEKPYECSKCGKAFSDGSALTQHQRIHTGEKPYECLECGKTFNRNSSLILHQRTHTGEKPYRCNECGKPFTDISHLTVHLRIHTGEKPYECSKCGKAFRDGSYLTQHERTHTGEKPFECAECGKSFNRNSHLIVHQKIHSGEKPYECKECGKTFIESAYLIRHQRIHTGEKPYGCNQCQKLFRNIAGLIRHQRTHTGEKPYECNQCGKAFRDSSCLTKHQRIHTKETPYQCPECGKSFKQNSHLAVHQRLHSREGPSRCPQCGKMFQKSSSLVRHQRAHLGEQPMET*WLGAT*VFQFTLTPVFRRRVLDLTPLWSVEKNPLSYPVN
3003	A	2	1489	SLTEHLSFFQPTAHSLTSLLGTMTTCSRQFTSSSSMKGSCGIGGGIGGGSSRISSVLAGGSCRAPSTYGGGLSVSSRFSSGGACGLGGGYGGGFSSSSSFGSGFGGGYGGGLGAGFGGGLGAGFGGGFAGGDGLLVGSEKVTMQNLNDRLASYLDKVRALEEANADLEVKIRDWYQRQRPSEIKDYSPYFKTIEDLRNKIIAATIENAQPILQIDNARLAADDFRTKYEHELALRQTVEADVNGLRRVLDELTLARTDLEMQIEGLKEELAYLRKNH*EEMLALRGQTGGEVNVETDAAPGVDLSCILNEMRNQYEQMAEKNRRDAETWFLSKT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EELNKEVASNSELVQSSRSEVTELRRVLQGLEIELQSQLSMKASLENSLEETKGRYCMQLSQIQGLIGSVEEQLAQLRCEMEQQSQEYQILLDVKTRLEQEIATYRRLLEGEDAHLSSQQASGQSYSSREVFTSSSSSSRQTRPILKEQSSSSFSQGQSS
3004	A	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGCTFKDKVLVAARRNASAVVLYNEERYGNITLPMSHAGTGNIVVIMISYPKGREILELVQKGIPVTMTIGVGTRHVQEFISGQSVVFVAIAFITMMIISLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVIGQLLLHTVKHGEKGIDVDAENCAVCIENFKVKDIIRILPCKHIFHRICIDPWLLDHRTCPMCKLDVIKALGYWGEPGDVQEMPAPESPPGRDPAANLSLALPDDDGSDESSPPSASPAESEPQCDPSFKGDAGENTALLEAGRSDSRHGGPIS
3005	A	184	2552	TMTIHQFLLLFLFWVCLPHFCSPEIMFRRTPVPQQRILSSRVPRSDGKILHRQKRGWMWNQFFLLEEYTGSDYQYVGKLHSDQDKGDGSLKYILSGDGAGTLFIIDEKTGDIHATRRIDREEKAFYTLRAQAINRRTLRPVEPESEFVIKIHDINDNEPTFPEEIYTASVPEMSVVGTSVVQVTATDADDPSYGNSARVIYSILQGQPYFSVEPETGIIRTALPNMNRENREQYQVVIQAKDMGGQMGGLSGTTTVNITLTDVNDNPPRFPQNTIHLRVLESSPVGTAIGSVKATDADTGKNAEVEYRIIDGDGTDMFDIVTEKDTQEGIITVKKPLDYESRRLYTLKVEAENTHVDPRFYYLGPFKDTTIVKISIEDVDEPPVFSRSSYLFEVHEDIEVGTIIGTVMARDPDSISSPIRFSLDRHTDLDRIFNIHSGNGSLYTSKPLDRELSQWHNLTVIAAEINNPKETTRVAVFVRILDANDNAPQFAVFYDTFVCENARPGQLIQTISAVDKDDPLGGQKFFFSLAAVNPNFTVQDNEDNTARILTRKNGFNRHEISTYLLPVVISDNDYPIQSSTGTLTIRVCACDSQGNMQSCSAEALLLPAGLSTGALIAILLCIIILLVIVVLFAALKRQRKKEPLILSKEDIRDNIVSYNDEGGGEEDTQAFDIGTLRNPAAIEEKKLRRDIIPETLFIPRRTPTAPDNTDVRDFINERLKEHDLDPTAPPYDSLATYAYEGNDSIAESLSSLESGTTEGDQNYDYLREWGPRFNKLPQKYGGGESDKDS
3006	A	2	541	GRVDKTWWGKSVGIMLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKKLLETECPQYIRKKGADVWFKELDINTDGAVNFQEFLILVIKMGVAALNSIIDVYHKYSLIKGNFHAVYRDDLQKLLETECPQYIRKKGADVWFKELDINTDGAVNFQEFLILVIKMGVGSPQKKVASYF
3007	A	1	1253	MYEGIRCLLKALLGFVSLAIGTLYCPRQYRPFPGSLGIEAINVPEPIPDSYYRDMATWPTHAPSVEEGGQGRFGNQADHFLGSLAFAKLLNRSLAVPSWIEYQHHKPPFTNLHVSYQKYFKLEPLQAYHRVISLEDFMEKLAPTHWPPEKRVAYCFEVAAQRSPDKKTCPMKEGNPFGPFWDQFHVSFNKSELFTGISFSASYREQWSQRFSPKEHPVLALPGAPAQFPVLEEHRPLQKYMVWSDEMVKTGEAQIHAHLVRPYVGIHLRIGSDWKNACAMLKDGTAGSHFMASPQCVGYSRSTAAPLTMTMCLPDLKEIQRAVKLWVRSLDAQSVYVATDSESYVPELQQLFKGKVKVVSLKPEVA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QVDLYILGQADHFIGNCVSSFTAFVKRERDLQGRPSSFFGMDRPPKLRDEF
3008	A	3136	1898	TARGGGSEPGPTMAANYSSTSTRREHVKVKTSSQPGFLERLSETSGGMFVGLMAFLLSFYLIFTNEGRALKTATSLAEGLSLVVSPDSIHSVAPENEGRLVHIIGALRTSKLLSDPNYGVHLPAVKLRRHVEMYQWVETEESREYTEDGQVKKETRYSYNTEWRSEIINSKNFDREIGHKNPRAMAGESFMATAPFVQIGRFFLSSGLIDKVDNFKSLSLSKLEDPHVDIIRRGDFFYHSENPKYPEVGDLRVSFSYAGLSGDDPDLGPAHVVTVIARQRGDQLVPFSTKSGDTLLLLHHGDFSAEEVFHRELRSNSMKTWGLRAAGWMAMFMGLNLMTRILYTLVDWFPVFRDLVNIGLKAFAFCVATSLTLLTVAAGWLFYRPLWALLIAGLALVPILVARTRVPAKKLE
3009	A	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPGGGSRGRSDRGSGQGDSLYPVGYLDKQVPDTSVQETDRILVEKRCWDIALGPLKQIPMNLFIMYMAGNTISIFPTMMVCMMAWRPIQALMAISATFKMLESSSQKFLQGLVYLIGNLMGLALAVYKCQSMGLLPTHASDWLAFIEPPERMEFSGGGLLL
3010	A	2	1041	LIDSAKARYWTQRGTWVYDNALLLLLKCLWSNVVPECTMASSNTVLMRLVASAYSIAQKAGMIVRRVIAEGDLGIVEKTCATDLQTKADRLAQMSICSSLARKFPKLTIIGEEDLPSEEVDQELIEDSQWEEILKQPCPSQYSAIKEEDLVVWVDPLDGTKEYTEGLLDNVTVLIGIAYEGKAIAGVINQPYYNYEAGPDAVLGRTIWGVLGLGAFGFQLKEVPAGKHIITTTRSHSNKLVTDCVAAMNPDAVLRVGGAGNKIIQLIEGKASAYVFASPGCKKWDTCAPEVILHAVGGKLTDIHGNVLQYHKDVKHMNSAGVLATLRNYDYYASRVPESIKNALVP
3011	A	291	1452	SPQKTMRSHTITMTTTSVSSWPYSSHRMRFITNHSDQPPQNFSATPNVTTCPMDEKLLSTVLTTSYSVIFIVGLVGNIIALYVFLGIHRKRNSIQIYLLNVAIADLLLIFCLPFRIMYHINQNKWTLGVILCKVVGTLFYMNMYISIILLGFISLDRYIKINRSIQQRKAITTKQSIYVCCIVWMLALGGFLTMIILTLKKGGHNSTMCFHYRDKHNAKGEAIFNFILVVMFWLIFLLIILSYIKIGKNLLRISKRRSKFPNSGKYATTARNSFIVLIIFTICFVPYHAFRFIYISSQLNVSSCYWKEIVHKTNEIMLVLSSFNSCLDPVMYFLMSSNIRKIMCQLLFRRFQGEPSRSESTSEFKPGYSLHDTSVAVKIQSSSKST
3012	A	246	1346	TEPVGYTKAEEPIAMRSLGALLLLLSACLAVSAGPVPTPPDNIQVQENFNISRIYGKWYNLAIGSTCPWLKKIMDRMTVSTLVLGEGATEAEISMTSTRWRKGVCEETSGAYEKTDTDGKFLYHKSKWNITMESYVVHTNYDEYAIFLTKKFSRHHGPTITAKLYGRAPQLRETLLQDFRVVAQGVGIPEDSIFTMADRGECVPGEQEPEPILIPRVRRAVLPQEEEGSGGGQLVTEVTKKEDSCQLGYSAGPCMGMTSRYFYNGTSMACETFQYGGCMGNGNNFVTEKECLQTCRTVAACNLPIVRGPCRAFIQLWAFDAVKGKCVLFPYGGCQGNGNKFYSEKECREYCGVPGDGDEELLRFSN
3013	A	67	379	RQMALLKANKDLISAGLKEFSVLLNQQVFNDPL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VSEEDMVTVVEDWMNFYINYYRQQVTGEPQERDKALQELRQELNTLANPFLAKYRDFLKSHELPSHPPPSS
3014	A	1	373	GTSWSTLRAVMSASVVSVVSRVLEEYLSSTPQRLKLLDAYLLYILLTGALQFGYCLFVLTFHFNSLLLFFFFCVGSFHSNVYFLLFTLSFLCFLFIAYFFLIRFFSLFIWFFHVFFIELSLFYF
3015	A	2	1321	AAAEGTAPSPGRVSPPTPARGEPEVTVEIGETYLCRRPDSTWHSAEVIQSRVNDQEGREEFYVHYVGFNRRLDEWVDKNRLALTKTVKDAVQKNSEKYLSELAEQPERKITRNQKRKHDEINHVQKTYAEMDPTTAALEKEHEAITKVKYVDKIHIGNYEIDAWYFSPFPEDYGKQPKLWLCEYCLKYMKYEKSYRFHLGQCQWRQPPGKEIYRKSNISVYEVDGKDHKIYCQNLCLLAKLFLDHKTLYFDVEPFVFYILTEVDRQGAHIVGYFSKEKESPDGNNVACILTLPPYQRRGYGKFLIAFSYELSKLESTVGSPEKPLSDLGKLSYRSYWSWVLLEILRDFRGTLSIKDLSQMTSITQNDIISTLQSLNMVKYWKGQHVICVTPKLVEEHLKSAQYKKPPITGGWGAAVCRGRWGSVSIWTGRSQGLLIAVT
3016	A	2	1321	AAAEGTAPSPGRVSPPTPARGEPEVTVEIGETYLCRRPDSTWHSAEVIQSRVNDQEGREEFYVHYVGFNRRLDEWVDKNRLALTKTVKDAVQKNSEKYLSELAEQPERKITRNQKRKHDEINHVQKTYAEMDPTTAALEKEHEAITKVKYVDKIHIGNYEIDAWYFSPFPEDYGKQPKLWLCEYCLKYMKYEKSYRFHLGQCQWRQPPGKEIYRKSNISVYEVDGKDHKIYCQNLCLLAKLFLDHKTLYFDVEPFVFYILTEVDRQGAHIVGYFSKEKESPDGNNVACILTLPPYQRRGYGKFLIAFSYELSKLESTVGSPEKPLSDLGKLSYRSYWSWVLLEILRDFRGTLSIKDLSQMTSITQNDIISTLQSLNMVKYWKGQHVICVTPKLVEEHLKSAQYKKPPITGGWGAAVCRGRWGSVSIWTGRSQGLLIAVT
3017	A	38	704	EAHPGGQLGSERNGVRMDEDVLTTLKILIIGESGVGKSSLLLRFTDDTFDPELAATIGVDFKVKTISVDGNKAKLAIWDTAGQERFRTLTPSYYRGAQGVILVYDVTRRDTFVKLDNWLNELETYCTRNDIVNMLVGNKIDKENREVDRNEGLKFARKHSMLFIEASAKTCDGVQCAFEELVEKIIQTPGLWESENQNKGVKLSHREEGQGGGACGGYCSVL
3018	A	2640	2861	APVLILQMVKLSIVLTPQFLSHDQGQLTKELQQHVKSVTCPCEYLRKVSECRQMGPGALEQFPGLSCHTSHSG
3019	A	1307	711	PGITMAASLVGKKIVFVTGNAKKLEEVVQILGDKFPCTLVAQKIDLPEYQGEPDEISIQKCQEAVRQVQGPVLVEDTCLCFNALGGLPGPYIKWFLEKLKPEGLHQLLAGFEDKSAYALCTFALSTGDPSQPVRLFRGRTSGRIVAPRGCQDFGWDPCFQPDGYEQTYAEMPKAEKNAVSHRFRALLELQEYFGSLAA
3020	A	1202	180	VSCLPTSCKMITLNNQDQPVPFNSSHPDEYKIAALVFYSCIFIIGLFVNITALWVFSCTTKKRTTVTIYMMNVALVDLIFIMTLPFRMFYYAKDEWPFGEYFCQILGALTVFYPSIALWLLAFISADRYMAIVQPKY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AKELKNTCKAVLACVGVWIMTLTTTTPLLLLYK DPDKDSTPATCLKISDIIYLKAVNVLNLTRLTFFF LIPLFIMIGCYLVIIHNLLHGRTSKLKPKVKEKSIRI IITLLVQVLVCFMPFHICFAFLMLGTGENSYNPW GAFTTFLMNLSTCLDVILYYIVSKQFQARVISVM LYRNYLRSMRRKSFRSGSLRSLSNINSEML
3021	A	27	1897	EEFCTWIAVRVGEMETAPKPGKDVPPKKDKLQT KRKKPRRYWEEETVPTTAGASPGPPRNKKNREL RPQRPKNAYILKKSRISKKPQVPKKPREWKNPES QRGLSGAQDPFPGPAPVPVEVVQKFCRIDKSRKL PHSKAKTRSRLEVAEAEEEETSIKAARSELLLAEE PGFLEGEDGEDTAKICQADIVEAVDIASAAKHFD LNLRQFGPYRLNYSRTGRHLAFGGRRGHVAALD WVTKKLMCEINVMEAVRDIRFLHSEALLAVAQN RWLHIYDNQGIELHCIRRCDRVTRLEFLPFHFLLA TASETGFLTYLDVSVGKIVAALNARAGRLDVMS QNPYNAVIHLGHSNGTVSLWSPAMKEPLAKILC HRGGVRAVAVDSTGTYMATSGLDHQLKIFDLRG TYQPLSTRTLPHGAGHLAFSQRGLLVAGMGDVV NIWAGQGKASPPSLEQPYLTHRLSGPVHGLQFCP FEDVLGVGHTGGITSMLVPGAGEPNFDGLESNPY RSRKQRQEWEVKALLEKVPAELICLDPRALAEV DVISLEQGKKEQIERLGYDPQAKAPFQPKPKQKG RSSTASLVKRKRKVMDEEHRDKVRQSLQQQHH KEAKAKPTGARPSALDRFVR
3022	A	1	2249	MTAQDSNTSAHAQRDGPELPASSSWRSFWPLSC LSSPPVSAVEVATEGRDREVAKVGQRFCDTTSGE LRQARDRDCCVRMPAPVGRRSPPSPRSSMAAVA LRDSAQGMTFEDVAIYFSQEEWELLDESQRFLYC DVMLENFAHVTSLGYCHGMENEAIASEQSVSIQ VRTSKGNTPTQKTHLSEIKMCVPVLKDILPAAEH QTTSPVQKSYLGSTSMRGFCFSADLHQHQKHYN EEEPWKRKVDEATFVTGCRFHVLNYFTCGEAFP APTDLLQHEATPSGEEPHSSSSKHIQAFFNAKSYY KWGEYRKASSHKHTLVQHQSVCSEGGLYECSK CEKAFTCKNTLVQHQQIHTGQKMFECSECEESFS KKCHLILHKIIHTGERPYECSDREKAFIHKSEFIHH QRRHTGGVRHECGECRKTFSYKSNLIEHQRVHT GERPYECGECGKSFRQSSSLFRHQRVHSGERPYQ CCECGKSFRQIFNLIRHRRVHTGEMPYQCSDCGK SFSCKSELIQHQRIHSGERPYECRECGKSFRQFSN LIRHRSIHTGDRPYECSECEKSFSRKFILIQHQRVH TGERPYECSECGKSFTRKSDLIQHRRIHTGTRPYE CSECGKSFRQRSGLIQHRRLHTGERPYECSECGK SFSQSASLIQHQRVHTGERPYQCCECGKSFRQIFN LIRHRRVHTGEMPYQCSDCGKSFSCKSELIQHRRI HSGERPYECSECGKSFSRKSNLIRHRRVHTEERP
3023	A	3148	634	AAGALRCLAAFPRAEPASRGRQSSPARACAASR AERATAAAMAHRCLRLWGRGGCWPRGLQQLL VPGGVGPGEQPCLRTLYRFVTTQARASRNSLLTD IIAAYQRFCSRPPKGFGKYFPNGKNGKKASEPKE VMGEKKESKPAATTRSSGGGGGGGGKRGGKKD DSHWWSRFQKGDIPWDDKDFRMFFLWTALFWG GVMFYLLLKRSGREITWKDFVNNYLSKGVVDRL EVVNKRFVRVTFTPGKTPVDGQYVWFNIGSVDT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				FERNLETLQQELGIEGENRVPVVYIAESDGSFLLSMLPTVLIIAFLLYTIRRGPAAIGRTGRGMGGLFSVGETTAKVLKDEIDVKFKDVAGCEEAKLEIMEFVNFLKNPKQYQDLGAKIPKGAILTGPPGTGKTLLAKATAGEANVPFITVSGSEFLEMFVGVGPARVRDLFALARKNAPCILFIDEIDAVGRKRGRGNFGGQSEQENTLNQLLVEMDGFNTTTNVVILAGTNRPDILDPALLRPGRFDRQIFIGPPDIKGRASIFKVHLRPLKLDSTLEKDKLARKLASLTPGFSGADVANVCNEAALIAARHLSDSINQKHFEQAIERVIGGLEKKTQVLQPEEKKTVAYHEAGHAVAGWYLEHADPLLKVSIIPRGKGLGYAQYLPKEQYLYTKEQLLDRMCMTLGGRVSEEIFFGRITTGAQDDLRKVTQSAYAQIVQFGMNEKVGQISFDLPRQGDMVLEKPYSEATARLIDDEVRILINDAYKRTVALLTEKKADVEKVALLLLEKEVLDKNDMVELLGPRPFAEKSTYEEFVEGTGSLDEDTSLPEGLKDWNKEREKEKEEPPGEKVAN
3024	A	274	1455	LRACSLPSMSALEKSMHLGRLPSRPPLPGSGGSQSGAKMRMGPGRKRDFSPVPWSQYFESMEDVEVENETGKDTFRVYKSGSEGPVLLLLHGGGHSALSWAVFTAAIISRVQCRIVALDLRSHGETKVKNPEDLSAETMAKDVGNVVEAMYGDLPPPIMLIGHSMGGAIAVHTASSNLVPSLLGLCMIDVVEGTAMDALNSMQNFLRGRPKTFKSLENAIEWSVKSGQIRNLESARVSMVGQVKQCEGITSPEGSKSIVEGIIEEEEEDEEGSESISKRKKEDDMETKKDHPYTWRIELAKTEKYWDGWFRGLSNLFLSCPIPKLLLLAGVDRLDKDLTIGQMQGKFQMQVLPQCGHAVHEDAPDKVAEAVATFLIRHRFAEPIGGFQCVFPGC
3025	A	621	306	YHGGQRGRAGGSFRSVQGWGGQLRNPFRTSKSLSWKGLSSLLFPLYNLQMGRPRDRKELGRGHSPPHLEGPHMLPSGAARWRWLEAPVLVLEPLVLRPAAAPTP
3026	A	1533	454	AKVPQSTREEKRENGLEARSPAINLMGFNVEEMYEAHAWIQRILSLQNHHIIENNHILYLGRKEHDILSQLQKTSSVSITEIISPGRTELEIEGARADLIEVVMNIEDMLCKVQEEMARKKERGLWRSLGQWTIQQQKTQDEMKENIIFLKCPVPPTQELLDQKKQFEKCGLQVLKVEKIDNEVLMAAFQRKKKMMEEKLHRQPVSHRLFQQVPYQFCNVVCRVGFQRMYSTPCDPKYGAGIYFTKNLKNLAEKAKKISAADKLIYVFEAEVLTGFFCQGHPLNIVPPPLSPGAIDGHDSVVDNVSSPETFVIFSGMQAIPQYLWTCTQEYVQSQDYSSGPMRPFAQHPWRGFASGSPVD
3027	A	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFIFSKSMNESMKNQKEFMLMNARLQLERQLIMQSEMRERQMAMQIAWSREFLKYFGTFFGLAAISLTAGAIKKKKPAFLVPIVPLSFILTYQYDLGYGTLLERMKGEAEDILETEKSKLQLPRGMITFESIEKARKEQSRFFIDK
3028	A	876	1226	AVGKEPESSSTWVRDREGHIRSRRSMKMLWKLTDNIKYEDCEVSATPARSSVRSQAPSLTLPLLLLSLQPAAKRGWDKLSPAQRPSLGFARRTRGRSCRERTWMLPSLVSEFLHRD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3029	A	3	1731	FREGRFGSSCAVAAPLAGFQGLIECGYLAVDSPPSCWTPGGSNPAAPLPQALLPPRLPPTVLPFLGPGLSGELEMFTLPQKDFRAPTTCLGPTCMQDLGSSHGEDLEGECSRKLDQKLPELRGVGDPAMISSNTSYLSSRGRMIKWFWDSAEEGYRTYHMDEYDEDKNPSGIINLGTSENKLCFDLLSWRLSQRDMQRVEPSLLQYADWRGHLFLREEVAKFLSFYCKSPVPLRPENVVVLNGGASLFSALATVLCEAGEAFLIPTPYYGAITQHVCLYGNIRLAYVYLDSEVTGLDTRPFQLTVEKLEMALREAHSEGVKVKGLILISPQNPLGDVYSPEELQEYLVFAKRHRLHVIVDEVYMLSVFEKSVGYRSVLSLERLPDPQRTHVMWATSKDFGMSGLRFGTLYTENQDVATAVASLCRYHGLSGLVQYQMAQLLRDRDWINQVYLPENHARLKAAHTYVSEELRALGIPFLSRGAGFFIWVDLRKYLLKGTFEEEMLLWRRFLDNKVLLSFGKAFECKEPGWFRFVFSDQVHRLCLGMQRVQQVLAGKSQVAEDPRPSQSQEPSDQRR
3030	A	1	584	PWLPWSDGRAARSSRKCPRSRFPVQVGKMAVSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFYRGDSPTDSQKDMIEIPLPPWQERTDESIETKRARLLYESRKRGMLENCILLSLFAKEHLQHMTEKQLNLYDRLINEPSNDWDIYYWATEAKPAPEIFENEVMALLRDFAKNKNKEQRLRAPDLEYLFEKPR
3031	A	1177	359	SLWPWILMDDSLMQISLQLLCVYTANFPNGCSSLCWSSCGQHPVQATHRGAVSNSLMLCILKLASQMPLENTTVQQMVFMLLSNLALSHDCKGVIQKSNFLQNFLSLALPKGGNKHLSNLTILWLKLLLNISSGEDGQQMILRLDGCLDLLTEMSKYKHKSSPLLPLLIFHNVCFSPANKPKILANEKVITVLAACLESENQNAQRIGAAALWALIYNYQKAKTALKSPSVKRRVDEAYSLAKKTFPNSEANPLNAYYLKCLENLVQLLNSS
3032	A	2	1242	GISGRPPRPAKRRMGKNPVRPPRALPPVPSQDDIPLSRPKKKKPRTKNTPASASLEGLAQTAGRRPSEGNEPSTKELKEHPEAPVQRRQKKTRLPLELETSSTQKKSSSSSLLRNENGIDAEPAEEAVIQKPRRKTKKTQPAELQYANELGVEDEDIITDEQTTVEQQSVFTAPTGISQPVGKVFVEKSRRFQAADRSELIKTTENIDVSMDVKPSWTTRDVALTVHRAFRMIGLFSHGFLAGCAVWNIVVIYVLAGDQLSNLSNLLQQYKTLAYPFQSLLYLLLALSTISAFDRIDFAKISVAIRNFLALDPTALASFLYFTALILSLSQQMTSDRIHLYTPSSVNGSLWEAGIEEQILQPWIVVNLVVALLVGLSWLFLSYRPGMDLSEELMFSSEVEEYPDKEKEIKASS
3033	A	3	1436	TATSGGIWLRRKWRCHWPRPLPQSCVGTEGGLQVRDTSSRIAKGGVDHTKMSLHGASGGHERSRDRRRSSDRSRDSSHERTESQLTPCIRNVTSPTRQHHVEREKDHSSSRPSSPRPQKASPNGSISSAGNSSRNSSQSSSDGSCKTAGEMVFVYENAKEGARNIRTSERVTLIVDNTRFVVDPSIFTAQPNTMLGRMFGSGREHNFTRPNEKGEYEVAEGIGSTVFRAILDYYKTGIIRCPDGISIPELREACDYLCISFEYSTIKCRDLSALMHELSNDGARRQFEFYLEEMILPLMVASAQSGERE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				CHIVVLTDDDVVDWDEEYPPQMGEEYSQIIYSTKLYRFFKYIENRDVAKSVLKERGLKKIRLGIEGYPTYKEKVKKRPGGRPEVIYNYVQRPFIRMSWEKEEGKSRHVDFQCVKSKSITNLAAAAADIPQDQLVVMHPTPQVDELDILPIHPPSGNSDLDPDAQNPML
3034	A	3	1972	SSLAQHRSVAVLGWPAGWAAARARPAMQGGNSGVRKREEEGDGAGAVAAPPAIDFPAEGPDPEYDESDVPAEIQVLKEPLQQPTFPFAVANQLLLVSLLEHLSHVHEPNPLRSRQVFKLLCQTFIKMGLLSSFTCSDEFSSLRLHHNRAITHLMRSAKERVRQDPCEDISRIQKIRSREVALEAQTSRYLNEFEELAILGKGGYGRVYKVRNKLDGQYYAIKKILIKGATKTVCMKVLREVKVLAGLQHPNIVGYHTAWIEHVHVIQPRADRAAIELPSLEVLSDQEEDREQCGVKNDESSSSSIIFAEPTPEKEKRFGESDTENQNNKSVKYTTNLVIRESGELESTLELQENGLAGLSASSIVEQQLPLRRNSHLEESFTSTEESSEENVNFLGQTEAQYHLMLHIQMQLCELSLWDWIVERNKRGREYVDESACPYVMANVATKIFQELVEGVFYIHNMGIVHRDLKPRNIFLHGPDQQVKIGDFGLACTDILQKNTDWTNRNGKRTPTHTSRVGTCLYASPEQLEGSEYDAKSDMYSLGVVLLELFQPFGTEMERAEVLTGLRTGQLPESLRKRCPVQAKYIQHLTRRNSSQRPSAIQLLQSELFQNSGNVNLTLQMKIIEQEKEIAELKKQLNLLSQDKGVRDDGKDGGVG
3035	A	110	1172	KLSCPCSHGTRVTAVRGPRLKAGVQWHDLGSLQPPPSGLKQSSHLSLSSSWDFRHAPTHPETYTCPKMIEMEQAEAQLAELDLLASMFPGENELIVNDQLAVAELKDCIEKKTMEGRSSKVYFTINMNLDVSDEKMAMFSLACILPFKYPAVLPEITVRSVLLSRSQQTQLNTDLTAFLQKHCHGDVCILNATEWVREHASGYVSRDTSSSPTTGSTVQSVDLIFTRLWIYSHHIYNKCKRKNILEWAKELSLSGFSMPGKPGVVCVEGPQSACEEFWARLRKLNWKRILIRHREDIPFDGTNDETERQRKFSIFEEKVFSVNGARGNHMDFGQLYQFLNTKGCGDVFQMFLWV
3036	A	1	2288	FRFAERRAAAAESDVSAKMAGRSMQAARCPTDELSLTNCAVVNEKDFQSGQHVIVRTSPNHRYTFTLKTHPSVVPGSIAFSLPQRKWAGLSIGQEIEVSLYTFDKAKQCIGTMTIEIDFLQKKSIDSNPYDTDKMAAEFIQQFNNQAFSVGQQLVFSFNEKLFGLLVKDIEAMDPSILNGEPATGKRQKIEVGLVVGNSQVAFEKAENSSLNLIGKAKTKENRQSIINPDWNFEKMGIGGLDKEFSDIFRRAFASRVFPPEIVEQMGCKHVKGILLYGPPGCGKTLLARQIGKMLNAREPKVVNGPEILNKYVGESEANIRKLFADAEEEQRRLGANSGLHIIIFDEIDAICKQRGSMAGSTGVHDTVVNQLLSKIDGVEQLNNILVIGMTNRPDLIDEALLRPGRLEVKMEIGLPDEKGRLQILHIHTARMRGHQLLSADVDIKELAVETKNFSGAELEGLVRAAQSTAMNRHIKASTKVEVDMEKAESLQVTRGDFLASLENDIKPAFGTNQEDYASYIMNGIIKWGDPVTRVLDDGELLVQQTKNSDRTPLVSVLLEGPPHSGKTALAAKIAEESNFPFIKICSPDKMIGFSETAKCQAMKKIFDDAYKSQLSCVVVDDIERLLDYVPIGPRFSNLVLQAL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LVLLKKAPPQGRKLLIIGTTSRKDVLQEMEMLNA FSTTIHVPNIATGEQLLEALELLGNFKDKERTTIA QQVKGKKVWIGIKKLLMLIEMSLQMDPEYRVRK FLALLREEGASPLDFD
3037	A	1	1347	MLDTGSEHLNRILKALPALQSAGSEGQNGSAESL GEGGTRDSDRARRKLRGGNKEIPTFYPCLVVRSP VTASDLRGTQDFAAYHGLSLILEPLGACNRLSVC VPVHSPPGMRVSPRSPSLRTLVIDPAEPAGAQRL RFSGKERSGEAGSAVEGLAVAVSMGDGGAERD RGPARRAESGGGGGRCGDRSGAGDLRADGGGH SPTEVAGTSASSPAGSRESGADSDGQPGPGEADH CRRILVRDAKGTIREIVLPKGLDLDRPKRTRTFFT AEQLYRLEMEFQRCQYVVGRERTELARQLNLSE TQVKVWFQNRRTKQKKDQSRDLEKRASSSASEA FATSNILRLLEQGRLLSVPRAPSLLALTPSLPGLP ASHRGTSLGDPRNSSPRLNPLSSASASPPLPPPLP AVCFSSAPLLDLPAGYELGSSAFEPYSWLERKVG SASSCKKANT
3038	A	924	501	TELLPLCSRSGPKPQSGDPLLQLAQQARPRLSGE RLETAPSLLLSRMACVISGWALSRGARTWTWAT PTGPVHRAQPAIRSLSAEGALTRLKEEKWPGRYI LPNHLTPPFLYKHLGSVPPSHWRSPLISHSVNILA LNWR
3039	A	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFLLIL ALGQAVQFQEYVFLQFLGLDKAPSPQKFQPVPYI LKKIFQDREAAATTGVSRDLCYVKELGVRGNVL RFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKE REQLTLAQLGLDLGPNSYYNLGPELELALFLVQE PHVWGQTTPKPGKMFVLRSVPWPQGAVHFNLL DVAKDWNDNPRKNFGLFLEILVKEDRDSGVNFQ PEDTCARLRCSLHASLLVVTLNPDQCHPSRKRRA AIPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAP KGFMANYCHGECPFSLTISLNSSNYAFMQALMH AVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHY EDMVVDECGCG
3040	A	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFALYL LSTRLPRGRRLGSTEEAGGRSLWFPSDLAELREL SEVLREYRKEHQAYVFLLFCGAYLYKQGFAIPGS SFLNVLAGALFGPWLGLLLCCVLTSVGATCCYL LSSIFGKQLVVSYFPDKVALLQRKVEENRNSLFF FLLFLRLFPMTPNWFLNLSAPILNIPIVQFFFSVLI GLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLL AIAMVALIPGTLIKKFSQKHLQLNETSTANHIHSR KDT
3041	A	1015	175	GLKRRRLCFAKVGDVLGCLSLPPSRSARVLEDISI LSCISVDSRIVRTKVPCSVTMSRPRKRLAGTSGSD KGLSGKRTKTENSGEALAKVEDSNPQKTSATKN CLKNLSSHWLMKSEPESRLEKGVDVKFSIEDLKA QPKQTTCWDGVRNYQARNFLRAMKLGEEAFFY HSNCKEPGIAGLMKIVKEAYPDHTQFEKNNPHY DPSSKEDNPKWSMVDVQFVRMMKRFIPLAELKS YHQAHKATGGPLKNMVLFTRQRLSIQPLTQEEF DFVLSLEEKEPS
3042	A	1015	175	GLKRRRLCFAKVGDVLGCLSLPPSRSARVLEDISI LSCISVDSRIVRTKVPCSVTMSRPRKRLAGTSGSD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KGLSGKRTKTENSGEALAKVEDSNPQKTSATKNCLKNLSSHWLMKSEPESRLEKGVDVKFSIEDLKAQPKQTTCWDGVRNYQARNFLRAMKLGEEAFFYHSNCKEPGIAGLMKIVKEAYPDHTQFEKNNPHYDPSSKEDNPKWSMVDVQFVRMMKRFIPLAELKSYHQAHKATGGPLKNMVLFTRQRLSIQPLTQEEFDFVLSLEEKEPS
3043	A	153	1133	VGTAPAPGGRDRAPAMGSFQLEDFAAGWIGGAASVIVGHPLDTVKTRLQAGVGYGNTLSCIRVVYRRESMFGFFKGMSFPLASIAVYNSVVFGVFSNTQRFLSQHRCGEPEASPPRTLSDLLLASMVAGVVSVGLGGPVDLIKIRLQMQTQPFRDANLGLKSRAVAPAEQPAYQGPVHCITTIVRNEGLAGLYRGASAMLLRDVPGYCLYFIPYVFLSEWITPEACTGPSPCAVWLAGGMAGAISWGTATPMDVVKSRLQADGVYLNKYKGVLDCISQSYQKEGLKVFFRGITVNAVRGFPMSAAMFLGYELSLQAIRGDHAVTSP
3044	A	41	1316	PPLGAGAGIHARSPHPARRLRLTAAGVGGRASGLLPTPWRRHHGPSGAAPYPAARLWQGPWRCRRPQPMAQRYDELPHYPGIADGPAALAGFPEAVPAAPGPYGPHRPPQPLPPGLDSDGLKRDKDEIYGHPLFPLLALGFEKCELATCSPRDGAGAGLGTPRGGDVCSSDSFNEDNTAFAKQVCSERPFSSNPELDNLMIQAIQVLRFHLLELEKGKMPIDLVIEDRDGGCREDFEDYPAPCPSLPDQNNIWIRDHEDSGSVHLGTPGPSSGGLASQSGDNSSDQGVGLDTSVASPSSGGEDEDLDQEPRRNKKRGIFPKVATNIMRAWLFQHLSHPYPSEEQKKQLAQDTGLTILQVNNWFINARRRIVQPMIDQSNRTGQGAAFSPEGQPIGGYTETEPHVAFRAPASVGMSLNSEGEWHYL
3045	A	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHACQVLILKHTHASLSLPSCQECFPSSIPSASHMVSHPHPPPSPRWGQTPEGLPAASPCGPGPRSCFSSILPTGDSWGMLACLCTVLWHLPAVPALNRTGDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLNYLGPPFNEPDFNPPRLGAETLPRATVDLEVWRSLNDKLRLTQNYEAYSHLLCYLRGLNRQAATAELRRSLAHFCTSLQGLLGSIAGVMAALGYPLPQPLPGTEPTWTPGPAHSDFLQKMDDFWLLKELQTWLWRSAKDFNRLKKKMQPPAAAVTLHLGAHGF
3046	A	1185	1584	MYAYMYICTHICICAYRGIHIDVYLYMCIYIHIWIHTYLCVHIYVYVYICTHICMCIHTYVYVYTYMYVYTYICLCVYICLCVHIYLCVYIHMYMCTHICMCIHTYVHMCICVYIHMYTCVYVYTYTCVYMY
3047	A	811	132	SLDLLGPIGILQEGRDPGTQGPQEKEKQMPASPMNTDAHLDINFKEGLKKERSYTGQFEANVRDEERQCGCGVVPDSLLMKVLSQRLDQQDCIQKGWVLHGVPRDLDQAHLLNRLGYNPNREFFLNVPFDSIMERLTLRRIDPVTGERYHLMYKPPPTMEIQARLLQNPKDAEEQVKLKMDLFYRNSADLEQLYGSAITLNGDQDPYTVFEYIESGIINPLPKKIP
3048	A	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRLHQYDGSIVVIQNPARQTLFFNGTRALKDERFQLEEFSPRRVRIRLSDARLEDEGGYFCQLYTEDTHHQIATLTVLVAPENPVVEVREQAVEGGEVELSCLVPRSR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PAATLRWYRDRKELKGVSSSQENGKVWSVAST VRFRVDRKDDGGIIICEAQNQALPSGHSKQTQYV LDVQYSPTARIHASQAVVREGDTLVLTCAVTGN PRPNQIRWNRGNESLPERAEAVGETLTLPGLVSA DNGTYTCEASNKHGHARALYVLVVYGESRLRPT EGGGGAPDPGAVVEAQTSVPYAIVGGILALLVFL IICVLVGMVWCSVRQKGSYLTHEASGLDEQGEA REAFLNGSDGHKRKEEFFI
3049	A	3159	882	VGCTLRVGVMAAAGSRKRRLAELTVDEFLASGF DSESESESENSPQAETREAREAARSPDKPGGSPSA SRRKGRASEHKDQLSRLKDRDPEFYKFLQENDQ SLLNFSDSDSSEEEEGPFHSLPDVLEEASEEEDGA EEGEDGDRVPRGLKGKKNSVPVTVAMVERWKQ AAKQRLTPKLFHEVVQAFRAAVATTRGDQESAE ANKFQVTDSAAFNALVTFCIRDLIGCLQKLLFGK VAKDSSRMLQPSSSPLWGKLRVDIKAYLGSAIQL VSCLSETTVLAAVLRHISVLVPCFLTFPKQCRML LKRMVVVWSTGEESLRVLAFLVLSRVCRHKKDT FLGPVLKQMYITYVRNCKFTSPGALPFISFMQWT LTELLALEPGVAYQHAFLYIRQLAIHLRNAMTTR KKETYQSVYNWQYVHCLFLWCRVLSTAGPSEA LQPLVYPLAQVIIGCIKLIPTARFYPLRMHCIRALT LLSGSSGAFIPVLPFILEMFQQVDFNRKPGRMSSK PINFSVILKLSNVNLQEKAYRDGLVEQLYDLTLE YLHSQAHCIGFPELVLPVVLQLKSFLRECKVANY CRQVQQLLGKVQENSAYICSRRQRVSFGVSEQQ AVEAWEKLTREEGTPLTLYYSHWRKLRDREIQL EISGKERLEDLNFPEIKRRKMADRKDEDRKQFKD LFDLNSSEEDDTEGFSERGILRPLSTRHGVEDDEE DEEEGEEDSSNSEDGDPDAEAGLAPGELQQLAQ GPEDELEDLQLSEDD
3050	A	870	182	HLDRYIKSPGSGSSTPAPPSHLLLYLLHPQSTRTM GCCGCSRGCGSGCGGCGSSCGGCGSGCGGCGSG RGGCGSGCGGCSSSCGGCGSRCYVPVCCCKPVC SWVPACSCTSCGSCGGSKGGCGSCGGSKGGCGS CGCSQSSCCKPCCCSSGCGSSCSQSSCCKPCCCSS GCGSSCCQSSCCKPYCCQSSCCKPCSCFSGCGSS CCQSSCYKPCCCQSSCCVPVCCQCKI
3051	A	175	4330	NIPRWNFQGKSFGVVLVHFSSEEVDMASDSPARS LDEIDLSALRDPAGIFELVELVGNGTYGQVYKGR HVKTGQLAAIKVMDVTGDEEEEIKQEINMLKKY SHHRNIATYYGAFIKKNPPGMDDQLWLVMEFCG AGSVTDLIKNTKGYTLKEEWIAYICREILRGLSHL HQHKVIHRDIKGQNVLLTENAEVKLVDFGVSAQ LDRTVGRRNTFIGTPYWMAPEVIACDENPDATY DFKSDLWSLGITAIEMAEGAPPLCDMHPMRALF LIPRNPAPRLKSKKWSKKFQSFIESCLVKNHSQRP ATEQLMKHPFIRDQPNERQVRIQLKDHIDRTKKK RGEKDETEYEYSGSEEEEEENDSGEPSSILNLPGE STLRRDFLRLQLANKERSEALRRQQLEQQQREN EEHKRQLLAERQKRIEEQKEQRRRLEEQQRREKE LRKQQEREQRRHYEEQMRREEERRRAEHEQEYI RRQLEEEQRQLEILQQQLLHEQALLLEYKRKQLE EQRQAERLQRQLKQERDYLVSLQHQRQEQRPVE KKPLYHYKEGMSPSEKPAWAKEVEERSRLNRQS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SPAMPHKVANRISDPNLPPRSESFSISGVQPARTP PMLRPVDPQIPHLVAVKSQGPALTASQSVHEQPT KGLSGFQEALNVTSHRVEMPRQNSDPTSENPPLP TRIEKFDRSSWLRQEEDIPPKVPQRTTSISPALAR KNSPGNGSALGPRLGSQPIRASNPDLRRTEPILES PLQRTSSGSSSSSSTPSSQPSSQGGSQPGSQAGSSE RTRVRANSKSEGSPVLPHEPAKVKPEESRDITRPS RPASYKKAIDEDLTALAKELRELRIEETNRPMKK VTDYSSSSEESESSEEEEEDGESETHDGTVAVSDI PRLIPTGAPGSNEQYNVGMVGTHGLETSHADSFS GSISREGTLMIRETSGEKKRSGHSDSNGFAGHINL PDLVQQSHSPAGTPTEGLGRVSTHSQEMDSGTE YGMGSSTKASFTPFVDPRVYQTSPTDEDEEDEES SAAALFTSELLRQEQAKLNEARKISVVNVNPTNI RPHSDTPEIRKYKKRFNSEILCAALWGVNLLVGT ENGLMLLDRSGQGKVYNLINRRRFQQMDVLEG LNVLVTISGKKNKLRVYYLSWLRNRILHNDPEV EKKQGWITVGDLEGCIHYKVVKYERIKFLVIALK NAVEIYAWAPKPYHKFMAFKSFADLQHKPLLVD LTVEEGQRLKVIFGSHTGFHVIDVDSGNSYDIYIP SHIQGNITPHAIVILPKTDGMEMLVCYEDEGVYV NTYGRITKDVVLQWGEMPTSVAYIHSNQIMGW GEKAIEIRSVETGHLDGVFMHKRAQRLKFLCERN DKVFFASVRSGGSSQVFFMTLNRNSMMNW
3052	A	1	615	MGQVECGGQKLGNQLEDDSEPAEGKVYSSDEE KLEASAGDPAGSEQEEEGSGGDSEDDGFLDSSA GGPGALLGPKPKLKGSLGTGAEEGAPVTAGVTA PGGKSRRRRTAFTSEQLLELEKEFHCKKYLSLTE RSQIAHALKLSEVQVKIWFQNRRAKWKRIKAGN VSSRSGEPVRNPKIVVPIPVHVNRFAVRSQHQQM EQGARP
3053	A	203	2167	FGVRVPSNTQCLVPSFHCMQTSEWDSECLTSLQP LPLPTPPAANEAHLQTAAISLWTVVAAVQAIERK VEIHSRRLLHLEGRTGTAEKKLASCEKTVTELGN QLEGKGAVLGTLLQEYGLLQRRLENLENLLRNR NFWILRLPPGIKGDIPKVPVAFDDVSIYFSTPEWE KLEEWQKELYKNIMKGNYESLISMDYAINQPDV LSQIQPEGEHNTEDQAGPEESEIPTDPSEEPGISTS DILSWIKQEEEPQVGAPPESKESDVYKSTYADEE LVIKAEGLARSSLCPEVPVPFSSPPAAAKDAFSDV AFKSQQSTSMTPFGRPATDLPEASEGQVTFTQLG SYPLPPPVGEQVFSCHHCGKNLSQDMLLTHQCS HATEHPLPCAQCPKHFTPQADLSSTSQDHASETP PTCPHCARTFTHPSRLTYHLRVHNSTERPFPCPDC PKRFADQARLTSHRRAHASERPFRCAQCGRSFSL KISLLLHQRGHAQERPFSCPQCGIDFNGHSALIRH QMIHTGERPYPCTDCSKSFMRKEHLLNHRRLHT GERPFSCPHCGKSFIRKHHLMKHQRIHTGERPYP CSYCGRSFRYKQTLKDHLRSGHNGGCGGDSDPS GQPPNPPGPLITGLETSGLGVNTEGLETNQWYGE GSGGGVL
3054	A	3	2212	SCGHKSAYGSYTGLQLFWEDGQELLQHQQLQD LRLCVHLRPQSEKVELSLWTLFVVGKGEPSAVR EKLGKAGFAAASGPGGRPGAERASTVLNILHLT AESRWEPNACNRVSSSPAGVGPLDLPVGPLLYFF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				APWARASFLCHAFQRPLTGIGLNTVRFTSEFPLH SKDPTAHKLLFTGNYLCKLHPRPRHAPQGSLSDF CHGTEGKDLPSEHNVSVEGVAQDRSPEATLCPQ KTCPCDICGLRLKDILHLAEHQTTHPRQKPFVCE AYVKGSEFSANLPRKQVQQNVHNPIRTEEGQAS PVKTCRDHTSDQLSTCREGGKDFVATAGFLQCE VTPSDGEPHEATEGVVDFHIALRHNKCCESGDAF NNKSTLVQHQRIHSRERPYECSKCGIFFTYAADL TQHQKVHNRGKPYECCECGKFFSQHSSLVKHRR VHTGESPHVCGDCGKFFSRSSNLIQHKRVHTGEK PYECSDCGKFFSQRSNLIHHKRVHTGRSAHECSE CGKSFNCNSSLIKHWRVHTGERPYKCNECGKFFS HIASLIQHQIVHTGERPHGCGECGKAFIRSSDLMK HQRVHTGERPYECNECGKLFSQSSSLNSHRRLHT GERPYQCSECGKFFNQSSSLNNHRRLHTGERPYE CSECGKTFRQRSNLRQHLKVHKPDRPYECSECG KAFNQRPTLIRHQKIHIRERSMENVLLPCSQHTPE ISSENRPYQGAVNYKLKLVHPSTHPGEVP
3055	A	268	2954	ARRSSSSQGSAAPTPCQVVEASRDQLVAGPSGK MGNREMEELIPLVNRLQDAFSALGQSCLLELPQI AVVGGQSAGKSSVLENFVGRDFLPRGSGIVTRRP LVLQLVTSKAEYAEFLHCKGKKFTDFDEVRLEIE AETDRVTGMNKGISSIPINLRVYSPHVLNLTLIDL PGITKVPVGDQPPDIEYQIRMIMQFITRENCLILA VTPANTDLANSDALKLAKEVDPQGLRTIGVITKL DLMDEGTDARDVLENKLLPLRRGYVGVVNRSQ KDIDGKKDIKAAMLAERKFFLSHPAYRHIADRM GTPHLQKVLNQQLTNHIRDTLPNFRNKLQGQLLS IEHEVEAYKNFKPEDPTRKTKALLQMVQQFAVD FEKRIEGSGDQVDTLELSGGAKINRIFHERFPFEIV KMEFNEKELRREISYAIKNIHGIRTGLFTPDMAFE AIVKKQIVKLKGPSLKSVDLVIQELINTVKKCTK KLANFPRLCEETERIVANHIREREGKTKDQVLLLI DIQVSYINTNHEDFIGFANAQQRSSQVHKKTTVG NQVIRKGWLTISNIGIMKGGSKGYWFVLTAESLS WYKDDEEKEKKYMLPLDNLKVRDVEKSFMSSK HIFALFNTEQRNVYKDYRFLELACDSQEDVDSW KASLLRAGVYPDKSVGNNKAENDENGQAENFS MDPQLERQVETIRNLVDSYMSIINKCIRDLIPKTI MHLMINNVKDFINSELLAQLYSSEDQNTLMEES AEQAQRRDEMLRMYQALKEALGIIGDIGTATVS TPAPPPVDDSWIQHSRRSPPPSPTTQRRPTLSAPL ARPTSGRGPAPAIPSPGPHSGAPPVPFRPGPLPPFP SSSDSFGAPPQVPSRPTRAPPSVPSRRPPPSPTRPTI IRPLESSLLD
3056	A	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVASGG WNDVACHTTMYFMCEFDKKNM
3057	A	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVASGG WNDVACHTTMYFMCEFDKKNM
3058	A	3363	2525	FLVKLILIILCRCLHSLSRSVQQLRTSFQDHAVWK PLMKVLQNAPDEILVVASSMLCNLLLEFSPSKEPI LESGAVELLCGLTQSENPALRVNGIWALMNMAF QAEQKIKADILRSLSTEQLFRLLSDSDLNVLMKT LGLLRNLLSTRPHIDKIMSTHGKQIMQAVTLILEG EHNIEVKEQTLCILANIADGTTAKDLIMTNDDILQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KIKYYMGHSHVKLQLAAMFCISNLIWNEEEGSQERQDKLRDMGIVDILHKLSQSPDSNLCDKAKMALQQYLA
3059	A	679	167	SSWPSLSSQMHFPSFHLHVAAHYGRDSFVRLLLEFKAEVDPLSDKGTTPLQLAIIRERSSCVKILLDHNANIDIQNGFLLRYAVIKSNHSYCRMFLQRGADTNLGRLEDGQTPLHLSALRDDVLCARMLYNYGADTNTRNYEGQTPLAVSISISGSSRPCLDFLQEVTSM
3060	A	30	234	PPLQLDMDPNCYCADGDSCTCAGSCKCKECKCTSCKKSCCSCCPAGCAKCAQGCICKGATDKCSCCA
3061	A	428	720	VRRDVRQQATWAMASDLDFSPPEVPEPTFLENLLRYGLFLGAIFQLICVLAIIVPIPKSHEAEAEPSEPRSAEVTRKPKAAVPSVNKRPKKETKKKR
3062	A	1589	276	WKQKYEPLGLDAAGIEEAITAVGSFILKANELLQVIDSSMKNFKAFFRWLYVAMLRMTEDHVLPELNKMTQKDITFVAEFLTEHFNEAPDLYNRKGKYFNVERVGQYLKDEDDDLVSPPNTEGNQWYDFLQNSSHLKESPLLFPYYPRKSLHFVKRRMENIIDQCLQKPADVIGKSMNQAICIPLYRDTRSEDSTRRLFKFPFLWNNKTSNLHYLLFTILEDSLYKMCILRRHTDISQSVSNGLIAIKFGSFTYATTEKVRRSIYSCLDAQFYDDETVTVVLKDTVGREGRDRLLVQLPLSLVYNSEDSAEYQFTGTYSTRLDEQCSAIPTRTMHFEKHWRLLESMKAQYVAGNGFRKVSCVLSSNLRHVRVFEMDIDDEWELDESSDEEEEASNKPVKIKEEVLSESEAENQQAGAAALAPEIVIKVEKLDPELDS
3063	A	50	849	DKMPSIFAYQSSEVDWCESNFQYSELVAEFYNTFSNIPFFIFGPLMMLLMHPYAQKRSRYIYVVWVLFMIIGLFSMYFHMTLSFLGQLLDEIAILWLLGSGYSIWMPRCYFPSFLGGNRSQFIRLVFITTVVSTLLSFLRPTVNAYALNSIALHILYIVCQEYRKTSNKELRHLIEVSVVLWAVALTSWISDRLLCSFWQRIHFFYLHSIWHVLISITFPYGMVTMALVDANYEMPGETLKVRYWPRDSWPVGLPYVEIRGDDKDC
3064	A	1523	925	AATMADGQMPFSCHYPSRLRRDPFRDSPLSSRLLDDGFGMDPFPDDLTASWPDWALPRLSSAWPGTLRSGMVPRGPTATARFGVPAEGRTPPPFPGEPWKVCVNVHSFKPEELMVKTKDGYVEVSGKHEEKQQEGGIVSKNFTKKIQLPAEVDPVTVFASLSPEGLLIIEAPQVPPYSTFGESSFNNELPQDSQEVTCT
3065	A	230	2929	LSTSLTGSHLFSLGNHSTRENLNAGNFNFPSEGHLVRSTGPGGSFAKHMVAQCVSPKGPLACSRTYFFGATHVPYLGGDSKLPKKTEQIRLLSQIYAAVIEAVLAGIACYAKTSSLTKAKEVAEQTLGSGLDSFELIPFKAALRSKMTFHIHAVNNQGRIVPLDSEDSLSFVKTACMAVYDIPDLLGGNGCLGSVVFSESFLTSQILVKEKDGTVTTETSSVVLTAAVPRFCSWLVEDNEVKLSEKTHQAVRGDESFLGTYLTGGEGAYLYSSNLQSWPEEGNVHFFSSGLLFSHCRHGSIIISKDHMNSISFYDGDSTSTVAALLIDFKSSLLPHLPVHFHGSSNFLMIALFPKSKIYQAFYSEVFSLWKQQDNSGISLKVIQEDGLSVEQKRLHSSAQKLFSALSQPAGEKRSSLKLLSAKLPELDWFLQHFAISSISQEPVMRTHLPVLLQQAEINTTHRIESDKVIISIVTGLPGCH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ASELCAFLVTLHKECGRWMVYRQIMDSSECFHAAHFQRYLSSALEAQQNRSARQSAYIRKKTRLLVVLQGYTDVIDVVQALQTHPDSNVKASFTIGAITACVEPMSCYMEHRFLFPKCLDQCSQGLVSNVVFTSHTTEQRHPLLVQLQSLIRAANPAAAFILAENGIVTRNEDIELILSENSFSSPEMLRSRYLMYPGWYEGKLNAGSVYPLMVQICVWFGRPLEKTRFVAKCKAIQSSIKPSPFSGNIYHILGKVKFSDSERTMEVCYNTLANSLSIMPVLEGPTPPPDSKSVSQDSSGQQECYLVFIGCSLKEDSIKDWLRQSAKQKPQRKALKTRGMLTQQEIRSIHVKRHLEPLPAGYFYNGTQFVNFFGDKTDFHPLMDQFMNDYVEEANREIEKYNQELEQQEYHDLFELKP
3066	A	130	588	LAPLRCQPGTRTQPRSHPAANDPSAAMSAAGARGLRATYHRLLDKVELMLPEKLRPLYNHPAGPRTVFFWAPIMKWGLVCAGLADMARPAEKLSTAQSAVLMATGFIWSRYSLVIIPKNWSLFAVNFFVGAAGASQLFRIWRYNQELKAKAHK
3067	A	2	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYPAGSLLRQSPQPRHTFYAGPRLSASASSKELLMKLRRKTGYSFVNCKKALETCGGDLKQAEIWLHKEAQKEGWSKAAKLQGRKTKEGLIGLLQEGNTTVLEVNCETDFVSRNLKFQLLVQQVALGTMMHCQTLKDQPSAYSKGFLNSSELSGLPAGPDREGSLKDQLALAIGKLGENMILKRAAWVKVPSGFYVGSYVHGAMQSPSLHKLVLGKYGALVICETSEQKTNLEDVGRRLGQHVVGMAPLSVGSLDDEPGGEAETKMLSQPYLLDPSITLGQYVQPQGVSVVDFVRFECGEGEEAAETE
3068	A	3	1679	NSRVWGPWTEPSAGSLRPMARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSGVPGGRRRKDWSCSLLVASLAGAFGSSFLYGYNLSVVNAPTPYIKAFYNESWERRHGRPIDPDTLTLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRKHTLLANNGFAISAALLMACSLQAGAFEMLIVGRFIMGIDGGVALSVLPMYLSEISPKEIRGSLGQVTAIFICIGVFTGQLLGLPELLGKESTWPYLFGVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEARAVKAFQTFLGKADVSQEVEEVLAESRVQRSIRLVSVLELLRAPYVRWQVVTVIVTMACYQLCGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFFGTLTITLTLQDHAPWVPYLSIVGILAIIASFCSGPGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSNFAVGLLFPFIQKSLDTYCFLVFATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTDGKINGRP
3069	A	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRLNRNARRKAAPRIECSHIRHAWDHAKSVRQNLAEMGLAVDPNRAVPLRKRKVKAMEVDIEERPKELVRKPYVLNDLEAEASLPEKKGNTLSRDLIDYVRYMVENHGEDYKAMARDEKNYYQDTPKQIRSKINVYKRFYPAEWQDFLDSLQKRKMEVE
3070	A	325	2019	LAEPEVATDSGQQADLPAEGGDPRAEASCSVLHSKPHAMADSRDPASDQMQHWKEQRAAQKADVLTTGAGNPVGDKLNVITVGPRGPLLVQDVVFTD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EMAHFDRERIPERVVHAKGAGAFGYFEVTHDIT KYSKAKVFEHIGKKTPIAVRFSTVAGESGSADTV RDPRGFAVKFYTEDGNWDLVGNNTPIFFIRDPILF PSFIHSQKRNPQTHLKDPDMVWDFWSLRPESLH QVSFLFSDRGIPDGHRHMNGYGSHTFKLVNANG EAVYCKFHYKTDQGIKNLSVEDAARLSQEDPDY GIRDLFNAIATGKYPSWTFYIQVMTFNQAETFPF NPFDLTKVWPHKDYPLIPVGKLVLNRNPVNYFA EVEQIAFDPSNMPPGIEASPDKMLQGRLFAYPDT HRHRLGPNYLHIPVNCPYRARVANYQRDGPMC MQDNQGGAPNYYPNSFGAPEQQPSALEHSIQYS GEVRRFNTANDDNVTQVRAFYVNVLNEEQRKR LCENIAGHLKDAQIFIQKKAVKNFTEVHPDYGSH IQALLDKYNAEKPKNAIHTFVQSGSHLAAREKA NL
3071	A	1	1187	SLGWLERPPALSRAAGDGARRLSGSRRGDVWLT SSAAGLLRSVAGGSWCGGQLRARGGSGRCVAR AMTGNAGEWCLMESDPGVFTELIKGFGCRGAQ VEEIWSLEPENFEKLKPVHGLIFLFKWQPGEEPA GSVVQDSRLDTIFFAKQVINNACATQAIVSVLLN CTHQDVHLGETLSEFKEFSQSFDAAMKGLALSN SDVIRQVHNSFARQQMFEFDTKTSAKEEDAFHF VSYVPVNGRLYELDGLREGPIDLGACNQDDWIS AVRPVIEKRIQKYSEGEIRFNLMAIVSDRKMIYEQ KIAELQRQLAEEEPMDTDQGNSMLSAIQSEVAK NQMLIEEEVQKLKRYKIENIRRKHNYLPFIMELL KTLAEHQQLIPLVEKAKEKQNAKKAQETK
3072	A	103	2775	RLRTLAPPGLLLGPPLVPDSRRRHQASLTPLHISG SPQLVGRGDRKLRTEVLVPPAALPAETRQRRSER LPRRTCPRGGAPGPGRSRLPRSLPPPSAIPGLRSPV WAAGLGGGGRREPSRGKGGAALRARHRSTMAE LGAGGDGHRGGDGAVRSETAPDSYKVQDKKNA SSRPASAISGQNNNHSGNKPDPPPVLRVDDRQRL ARERREEREKQLAAREIVWLEREERARQHYEKH LEERKKRLEEQRQKEERRRAAVEEKRRQRLEED KERHEAVVRRTMERSQKPKQKHNRWSWGGSLH GSPSIHSADPDRRSVSTMNLSKYVDPVISKRLSSS SATLLNSPDRARRLQLSPWESSVVNRLLTPTHSF LARSKSTAALSGEAVIPICPRSASCSPIIMPYKAAH SRNSMDRPKLFVTPPEGSSRRRIIHGTASYKKERE RENVLFLTSGTRRAVSPSNPKARQPARSRLWLPS KSLPHLPGTPRPTSSLPPGSVKAAPAQVRPPSPGN IRPVKREVKVEPEKKDPEKEPQKVANEPSLKGRA PLVKVEEATVEERTPAEPEVGPAAPAMAPAPAS APAPASAPAPAPVPTPAMVSAPSSTVNASASVKT SAGTTDPEEATRLLAEKRRLAREQREKEERERRE QEELERQKREELAQRVAEERTTRREEESRRLEAE QAREKEEQLQRQAEERALREWEEAERAQRQKEE EARVREEAERVRQEREKHFQREEQERLERKKRL EEIMKRTRRTEATDKKTSDQRNGDIAKGALTGG TEVSALPCTTNAPGNGKPVGSPHVVTSHQSKVT VESTPDLEKQPNENGVSVQNENFEEIINLPIGSKP SRLDVTNSESPEIPLNPILAFDDEGTLGPLPQVDG VQTQQTAEVI
3073	A	67	2415	PPRVCRDHVCLICWDPIAGTGGSRSTMPALPLDQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LQITHKDPKTGKLRTSPALHPEQKADRYFVLYKP PPKDNIPALVEEYLERATFVANDLDWLLALPHD KFWCQVIFDETLQKCLDSYLRYVPRKFDEGVAS APEVVDMQKRLHRSVFLTFLRMSTHKESKDHFIS PSAFGEILYNNFLFDIPKILDLCVLFGKGNSPLLQ KMIGNIFTQQPSYYSDLDETLPTILQVFSNILQHC GLQGDGANTTPQKLEERGRLTPSDMPLLELKDIV LYLCDTCTTLWAFLDIFPLACQTFQKHDFCYRLA SFYEAAIPEMESAIKKRRLEDSKLLGDLWQRLSH SRKKLMEIFHIILNQICLLPILESSCDNIQGFIEEFL QIFSSLLQEKRFLRDYDALFPVAEDISLLQQASSV LDETRTAYILQAVESAWEGVDRRKATDAKDPSV IEEPNGEPNGVTVTAEAVSQASSHPENSEEEECM GAAAAVGPAMCGVELDSLISQVKDLLPDLGEGFI LACLEYYHYDPEQVINNILEERLAPTLSQLDRNL DREMKPDPTPLLTSRHNVFQNDEFDVFSRDSVDL SRVHKGKSTRKEENTRSLLNDKRAVAAQRQRYE QYSVVVEEVPLQPGESLPYHSVYYEDEYDDTYD GNQVGANDADSDDELISRRPFTIPQVLRTKVPRE GQEEDDDDEEDDADEEAPKPDHFVQDPAVLREK AEARRMAFLAKKGYRHDSSTAVAGSPRGHGQS RETTQERRKKEANKATRANHNRRTMADRKRSK GMIPS
3074	A	3	251	GEARSPPPAAALLDMDPETCPCPSGGSCTCADSC KCEGCKCTSCKKSCCSCCPAECEKCAKDCVCKG GEAAEAEAEKCSCCQ
3075	A	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQRL RKFRELHLMRNEARKLNHQEVVEEDKRLKLPAN WEAKKARLEWELKEEEKKKECAARGEDYEKVK LLEISAEDAERWERKKKRKNPDLGFSDYAAAQL RQYHRLTKQIKPDMETYERLREKHGEEFFPTSNS LLHGTHVPSTEEIDRMVIDLEKQIEKRDKYSRRR PYNDDADIDYINERNAKFNKKAERFYGKYTAEI KQNLERGTAV
3076	A	255	982	SQFSLSQVLVDSAEEGSLAAAAELAAQKREQRL RKFRELHLMRNEARKLNHQEVVEEDKRLKLPAN WEAKKARLEWELKEEEKKKECAARGEDYEKVK LLEISAEDAERWERKKKRKNPDLGFSDYAAAQL RQYHRLTKQIKPDMETYERLREKHGEEFFPTSNS LLHGTHVPSTEEIDRMVIDLEKQIEKRDKYSRRR PYNDDADIDYINERNAKFNKKAERFYGKYTAEI KQNLERGTAV
3077	A	1	968	FRLRPRRACAQLLWHPAAGMASWAKGRSYLAP GLLQGQVAIVTGGATGIGKAIVKELLELGSNVVI ASRKLERLKSAADELQANLPPTKQARVIPIQCNIR NEEEVNNLVKSTLDTFGKINFLVNNGGGQFLSPA EHISSKGWHAVLETNLTGTFYMCKAVYSSWMK KHGGSIVNIIVPTKAGFPLAVHSGAARAGVYNLT KSLAFEWACSGIRINCVAPGVIYSQTAVENYGSW GQSFFEGSFQKIPAKRIGVPEEVSSVVCFLLSPAA SFITGQSVDVDGGRSLYTHSYEVPDHDNWPKGA GDLSVVKKMKETFKEKAKL
3078	A	2	3508	FVRESGKAPVTFDDITVYLLQEEWVLLSQQQKEL CGSNKLVAPLGPTVANPELFRKFGRGPEPWLGS VQGQRSLLEHHPGKKQMGYMGEMEVQGPTRES

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GQSLPPQKKAYLSHLSTGSGHIEGDWAGRNRKL LKPRSIQKSWFVQFPWLIMNEEQTALFCSACREY PSIRDKRSRLIEGYTGPFKVETLKYHAKSKAHMF CVNALAARDPIWAARFRSIRDPPGDVLASPEPLF TADCPIFYPPGPLGGFDSMAELLPSSRAELEDPGG DGAIPAMYLDCISDLRQKEITDGIHSSSDINILYN DAVESCIQDPSAEGLSEEVPVVFEELPVVFEDVA VYFTREEWGMLDKRQKELYRDVMRMNYELLAS LGPAAAKPDLISKLERRAAPWIKDPNGPKWGKG RPPGNKKMVAVREADTQASAADSALLPGSPVEA RASCCSSSICEEGDGPRRIKRTYRPRSIQRSWFGQ FPWLVIDPKETKLFCSACIERPNLHDKSSRLVRG YTGPFKVETLKYHEVSKAHRLCVNTVEIKEDTPH TALVPEISSDLMANMEHFFNAAYSIAYHSRPLND FEKILQLLQSTGTVILGKYRNRTACTQFIKYISETL KREILEDVRNSPCVSVLLDSSTDASEQACVGIYIR YFKQMEVKESYITLAPLYSETADGYFETIVSALD ELDIPFRKPGWVVGLGTDGSAMLSCRGGLVEKF QEVIPQLLPVHCVAHRLHLAVVDACGSIDLVKK CDRHIRTVFKFYQSSNKRLNELQEGAAPLEQEIIR LKDLNAVRWVASRRRTLHALLVSWPALARHLQ RVAEAGGQIGHRAKGMLKLMRGFHFVKFCHFL LDFLSIYRPLSEVCQKEIVLITEVNATLGRAYVAL ESLRHQAGPKEEEFNASFKDGRLHGICLDKLEVA EQRFQADRERTVLTGIEYLQQRFDADRPPQLKN MEVFDTMAWPSGIELASFGNDDILNLARYFECSL PTGYSEEALLEEWLGLKTIAQHLPFSMLCKNALA QHCRFPLLSKLMAVVVCVPISTSCCERGFKAMN RIRTDERTKLSNEVLNMLMMTAVNGVAVTEYD PQPAIQHWYLTSSGRRFSHVYTCAQVPARSPASA RLRKEEMGALYVEEPRTQKPPILPSREAAEVLKD CIMEPPERLLYPHTSQEAPGMS
3079	A	343	1513	FSPLEPRLCSLGGWGALQAGEPCQPSRAGCGRE GATMGCTLSAEERAALERSKAIEKNLKEDGISAA KDVKLLLLGAGESGKSTIVKQMKIIHEDGFSGED VKQYKPVVYSNTIQSLAAIVRAMDTLGIEYGDK ERKADAKMVCDVVSRMEDTEPFSAELLSAMMR LWGDSGIQECFNRSREYQLNDSAKYYLDSLDRIG AADYQPTEQDILRTRVKTTGIVETHFTFKNLHFR LFDVGGQRSERKKWIHCFEDVTAIIFCVALSGYD QVLHEDETTNRMHESLKLFDSICNNKWFTDTSII LFLNKKDIFEEKIKKSPLTICFPEYTGPSAFTEAVA YIQAQYESKNKSAHKEIYSHVTCATDTNNIQFVF DAVTDVIIAKNLRGCGLY
3080	A	41	997	EARTARELTDGVTDGLTMADQPKPISPLKNLLA GGFGGVCLVFVGHPLDTVKVRLQTQPPSLPGQPP MYSGTFDCFRKTLFREGITGLYRGMAAPIIGVTP MFAVCFFGFGLGKKLQQKHPEDVLSYPQLFAAG MLSGVFTTGIMTPGERIKCLLQIQASSGESKYTGT LDCAKKLYQEFGIRGIYKGTVLTLMRDVPASGM YFMTYEWLKNIFTPEGKRVSELSAPRILVAGGIA GIFNWAVAIPPDVLKSRFQTAPPGKYPNGFRDVL RELIRDEGVTSLYKGFNAVMIRAFPANAACFLGF EVAMKFLNWATPNL
3081	A	3	1996	IMADMEDLFGSDADSEAERKDSDSGSDSDSDQE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NAASGSNASGSESDQDERGDSGQPSNKELFGDD SEDEGASHHSGSDNHSERSDNRSEASERSDHEDN DPSDVDQHSGSEAPNDDEDEGHRSDGGSHHSEA EGSEKAHSDDEKWGREDKSDQSDDEKIQNSDDE ERAQGSDEDKLQNSDDDEKMQNTDDEERPQLS DDERQQLSEEEKANSDDERPVASDNDDEKQNSD DEEQPQLSDEEKMQNSDDERPQASDEEHRHSDD EEEQDHKSESARGSDSEDEVLRMKRKNAIASDSE ADSDTEVPKDNSGTMDLFGGADDISSGSDGEDK PPTPGQPVDENGLPQDQQEEEPIPETRIEVEIPKV NTDLGNDLYFVKLPNFLSVEPRPFDPQYYEDEFE DEEMLDEEGRTRLKLKVENTIRWRIRRDEEGNEI KESNARIVKWSDGSMSLHLGNEVFDVYKAPLQG DHNHLFIRQGTGLQGQAVFKTKLTFRPHSTDSAT HRKMTLSLADRCSKTQKIRILPMAGRDPECQRTE MIKKEEERLRASIRRESQQRRMREKQHQRGLSAS YLEPDRYDEEEEGEESISLAAIKNRYKGGIREERA RIYSSDSDEGSEEDKAQRLLKAKKLTSDEVRPNL FNSRGLSCTQEPTALNEELTDQAGTN
3082	A	3	921	VEFCLPASADSSSLVAASLAGVRKMATNFLAHE KIWFDKFKYDDAERRFYEQMNGPVAGASRQEN GASVILRDIARARENIQKSLAGSSGPGASSGTSGD HGELVVRIASLEVENQSLRGVVQELQQAISKLEA RLNVLEKSSPGHRATAPQTQHVSPMRQVEPPAK KPATPAEDDEDDDIDLFGSDNEEEDKEAAQLREE RLRQYAEKKAKKPALVAKSSILLDVKPWDDETD MAQLEACVRSIQLDGLVWGASKLVPVGYGIRKL QIQCVVEDDKVGTDLLEEEITKFEEHVQSVDIAA FNKI
3083	A	3	921	VEFCLPASADSSSLVAASLAGVRKMATNFLAHE KIWFDKFKYDDAERRFYEQMNGPVAGASRQEN GASVILRDIARARENIQKSLAGSSGPGASSGTSGD HGELVVRIASLEVENQSLRGVVQELQQAISKLEA RLNVLEKSSPGHRATAPQTQHVSPMRQVEPPAK KPATPAEDDEDDDIDLFGSDNEEEDKEAAQLREE RLRQYAEKKAKKPALVAKSSILLDVKPWDDETD MAQLEACVRSIQLDGLVWGASKLVPVGYGIRKL QIQCVVEDDKVGTDLLEEEITKFEEHVQSVDIAA FNKI
3084	A	128	4050	KSIVKIRKRMAAETQTLNFGPEWLRALSSGGSITS PPLSPALPKYKLADYRYGREEMLALFLKDNKIPS DLLDKEFLPILQEEPLPPLALVPFTEEEQRNFSMS VNSAAVLRLTGRGGGGTVVGAPRGRSSSRGRGR GRGECGFYQRSFDEVEGVFGRGGGREMHRSQS WEERGDRRFEKPGRKDVGRPNFEEGGPTSVGRK HEFIRSESENWRIFREEQNGEDEDGGWRLAGSRR DGERWRPHSPDGPRSAGWREHMERRRRFEFDFR DRDDERGYRRVRSGSGSIDDDRDSLPEWCLEDA EEEMGTFDSSGAFLSLKKVQKEPIPEEQEMDFRP VDEGEECSDSEGSHNEEAKEPDKTNKKEGEKTD RVGVEASEETPQTSSSSARPGTPSDHQSQEASQFE RKDEPKTEQTEKAEEETRMENSLPAKVPSRGDE MVADVQQPLSQIPSDTASPLLILPPPVPNPSPTLRP VETPVVGAPGMGSVSTEPDDEEGLKHLEQQAEK MVAYLQDSALDDERLASKLQEHRAKGVSIPLMH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EAMQKWYYKDPQGEIQGPFNNQEMAEWFQAG YFTMSLLVKRACDESFQPLGDIMKMWGRVPFSP GPAPPPHMGELDQERLTRQQELTALYQMQHLQY QQFLIQQQYAQVLAQQQKAALSSQQQQQLALLL QQFQTLKMRISDQNIIPSVTRSVSVPDTGSIWELQ PTASQPTVWEGGSVWDLPLDTTTPGPALEQLQQ LEKAKAAKLEQERREAEMRAKREEEERKRQEEL RRRQKGILRRQQEEERKRREEEELARRKQEEALR RQREQEIALRRQREEEERQQQEEALRRLEERRRE EEERRKQEELLRKQEEEAAKWAREEEEAQRRLE ENRLRMEEEAARLRHEEEERKRKELEVQRQKEL MRQRQQQQEALRRLQQQQQQQLAQMKLPSSS TWGQQSNTTACQSQATLSLAEIQKLEEERERQLR EEQRRQQRELMKALQQQQQQQQKLSGWGNV SKPSGTTKSLLEIQQEEARQMQKQQQQQQHQQ PNRARNNTHSNLHTSIGNSVWGSINTGPPNQWA SDLVSSIWSNADTKNSNMGFWDDAVKEVGPRN STNKNKNNASLSKSVGVSNRQNKKVEEEEKLLK LFQGVNKAQDGFTQWCEQMLHALNTANNLDVP TFVSFLKEVESPYEVHDYIRAYLGDTSEAKEFAK QFLERRAKQKANQQRQQQQLPQQQQQPPQQPP QQPQQQDSVWGMNHSTLHSVFQTNQSNNQQSN FEAVQSGKKKKKQKMVRADPSLLGFSVNASSER LNMGEIETLDDY
3085	A	128	4050	KSIVKIRKRMAAETQTLNFGPEWLRALSSGGSITS PPLSPALPKYKLADYRYGREEMLALFLKDNKIPS DLLDKEFLPILQEEPLPPLALVPFTEEEQRNFSMS VNSAAVLRLTGRGGGGTVVGAPRGRSSSRGRGR GRGECGFYQRSFDEVEGVFGRGGGREMHRSQS WEERGDRRFEKPGRKDVGRPNFEEGGPTSVGRK HEFIRSESENWRIFREEQNGEDEDGGWRLAGSRR DGERWRPHSPDGPRSAGWREHMERRRRFEFDFR DRDDERGYRRVRSGSGSIDDDRDSLPEWCLEDA EEEMGTFDSSGAFLSLKKVQKEPIPEEQEMDFRP VDEGEECSDSEGSHNEEAKEPDKTNKKEGEKTD RVGVEASEETPQTSSSSARPGTPSDHQSQEASQFE RKDEPKTEQTEKAEEETRMENSLPAKVPSRGDE MVADVQQPLSQIPSDTASPLLILPPPVPNPSPTLRP VETPVVGAPGMGSVSTEPDDEEGLKHLEQQAEK MVAYLQDSALDDERLASKLQEHRAKGVSIPLMH EAMQKWYYKDPQGEIQGPFNNQEMAEWFQAG YFTMSLLVKRACDESFQPLGDIMKMWGRVPFSP GPAPPPHMGELDQERLTRQQELTALYQMQHLQY QQFLIQQQYAQVLAQQQKAALSSQQQQQLALLL QQFQTLKMRISDQNIIPSVTRSVSVPDTGSIWELQ PTASQPTVWEGGSVWDLPLDTTTPGPALEQLQQ LEKAKAAKLEQERREAEMRAKREEEERKRQEEL RRRQKGILRRQQEEERKRREEEELARRKQEEALR RQREQEIALRRQREEEERQQQEEALRRLEERRRE EEERRKQEELLRKQEEEAAKWAREEEEAQRRLE ENRLRMEEEAARLRHEEEERKRKELEVQRQKEL MRQRQQQQEALRRLQQQQQQQLAQMKLPSSS TWGQQSNTTACQSQATLSLAEIQKLEEERERQLR EEQRRQQRELMKALQQQQQQQQKLSGWGNV SKPSGTTKSLLEIQQEEARQMQKQQQQQQHQQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PNRARNNTHSNLHTSIGNSVWGSINTGPPNQWA SDLVSSIWSNADTKNSNMGFWDDAVKEVGPRN STNKNKNNASLSKSVGVSNRQNKKVEEEEKLLK LFQGVNKAQDGFTQWCEQMLHALNTANNLDVP TFVSFLKEVESPYEVHDYIRAYLGDTSEAKEFAK QFLERRAKQKANQQRQQQQLPQQQQQPPQQPP QQPQQQDSVWGMNHSTLHSVFQTNQSNNQQSN FEAVQSGKKKKKQKMVRADPSLLGFSVNASSER LNMGEIETLDDY
3086	A	675	1334	LHPAATSTAWLHVPPGLSMALSWVLTVLSLLPL LEAQIPLCANLVPVPITNATLDRITGKWFYIASAF RNEEYNKSVQEIQATFFYFTPNKTEDTIFLREYQT RQDQCIYNTTYLNVQRENGTISRYVGGQEHFAH LLILRDTKTYMLAFDVNDEKNWGLSVYADKPET TKEQLGEFYEALDCLRIPKSDVVYTDWKKDKCE PLEKQHEKERKQEEGES
3087	A	1	1575	CTPVARSMATTATCTRFTDDYQLFEELGKGAFS VVRRCVKKTSTQEYAAKIINTKKLSARDHQKLE REARICRLLKHPNIVRLHDSISEEGFHYLVFDLVT GGELFEDIVAREYYSEADASHCIHQILESVNHIHQ HDIVHRDLKPENLLLASKCKGAAVKLADFGLAIE VQGEQQAWFGFAGTPGYLSPEVLRKDPYGKPVD IWACGVILYILLVGYPPFWDEDQHKLYQQIKAG AYDFPSPEWDTVTPEAKNLINQMLTINPAKRITA DQALKHPWVCQRSTVASMMHRQETVECLRKFN ARRKLKGAILTTMLVSRNFSAAKSLLNKKSDGG VKPQSNNKNSLVSPAQEPAPLQTAMEPQTTVVH NATDGIKGSTESCNTTTEDEDLKVRKQEIIKITEQ LIEAINNGDFEAYTKICDPGLTSFEPEALGNLVEG MDFHKFYFENLLSKNSKPIHTTILNPHVHVIGED AACIAYIRLTQYIDGQGRPRTSQSEETRVWHRRD GKWLNVHYHCSGAPAAPLQ
3088	A	12	1039	SSVAEFPERVQLSQPQNWNFSGAGGAWSLDFAE QLKWSAELARLGESIMDGKQGGMDGSKPAGPR DFPGIRLLSNPLMGDAVSDWSPMHEAAIHGHQL SLRNLISQGWAVNIITADHVSPLHEACLGGHLSC VKILLKHGAQVNGVTADWHTPLFNACVSGSWD CVNLLLQHGASVQPESDLASPIHEAARRGHVEC VNSLIAYGGNIDHKISHLGTPLYLACENQQRACV KKLLESGADVNQGKGQDSPLHAVARTASEELAC LLMDFGADTQAKNAEGKRPVELVPPESPLAQLF LEREGPPSLMQLCRLRIRKCFGIQQHHKITKLVLP EDLKQFLLHL
3089	A	73	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPSPCC MFFVSKRIPENRVVSYQLSSRSTCLKAGVIFTTKK GQQFCGDPKQEWVQRYMKNLDAKQKKASPRA RAVAVKGPVQRYPGNQTTC
3090	A	4627	611	LMEAGGGGGALPAGVETMVLTLGESWPVLVGR RFLSLSAADGSDGSHDSWDVERVAEWPWLSGTI RAVSHTDVTKKDLKVCVEFDGESWRKRRWIEV YSLLRRAFLVEHNLVLAERKSPEISERIVQWPAIT YKPLLDKAGLGSITSVRFLGDQQRVFLSKDLLKP IQDVNSLRLSLTDNQIVSKEFQALIVKHLDESHLL KGDKNLVGSEVKIYSLDPSTQWFSATVVNGNPA SKTLQVNCEEIPALKIVDPSLIHVEVVHDNLVTC

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GNSARIGAVKRKSSENNGTLVSKQAKSCSEASPS MCPVQSVPTTVFKEILLGCTAATPPSKDPRQQST PQAANSPPNLGAKIPQGCHKQSLPEEISSCLNTKS EALRTKPDVCKAGLLSKSSQIGTGDLKILTEPKGS CTQPKTNTDQENRLESVPQALTGLPKECLPTKAS SKAELEIANPPELQKHLEHAPSPSDVSNAPEVKA GVNSDSPNNCSGKKVEPSALACRSQNLKESSVK VDNESCCSRSNNKIQNAPSRKSVLTDPAKLKKLQ QSGEAFVQDDSCVNIVAQLPKCRECRLDSLRKD KEQQKDSPVFCRFFHFRRLQFNKHGVLRVEGFLT PNKYDNEAIGLWLPLTKNVVGIDLDTAKYILANI GDHFCQMVISEKEAMSTIEPHRQVAWKRAVKG VREMCDVCDTTIFNLHWVCPRCGFGVCVDCYR MKRKNCQQGAAYKTFSWLKCVKSQIHEPENLM PTQIIPGKALYDVGDIVHSVRAKWGIKANCPCSN RQFKLFSKPASKEDLKQTSLAGEKPTLGAVLQQ NPSVLEPAAVGGEAASKPAGSMKPACPASTSPLN WLADLTSGNVNKENKEKQPTMPILKNEIKCLPPL PPLSKSSTVLHTFNSTILTPVSNNNSGFLRNLLNSS TGKTENGLKNTPKILDDIFASLVQNKTTSDLSKR PQGLTIKPSILGFDTPHYWLCDNRLLCLQDPNNK SNWNVFRECWKQGQPVMVSGVHHKLNSELWK PESFRKEFGEQEVDLVNCRTNEIITGATVGDFWD GFEDVPNRLKNEKEPMVLKLKDWPPGEDFRDM MPSRFDDLMANIPLPEYTRRDGKLNLASRLPNYF VRPDLGPKMYNAYGLITPEDRKYGTTNLHLDVS DAANVMVYVGIPKGQCEQEEEVLKTIQDGDSDE LTIKRFIEGKEKPGALWHIYAAKDTEKIREFLKK VSEEQGQENPADHDPIHDQSWYLDRSLRKRLHQ EYGVQGWAIVQFLGDVVFIPAGAPHQVHNLYSC IKVAEDFVSPEHVKHCFWLTQEFRYLSQTHTNHE DKLQVKNVIYHAVKDAVAMLKASESSFGKP
3091	A	97	1838	KRGARRGGWKRKMPSTDLLMLKAFEPYLEILEV YSTKAKNYVNGHCTKYEPWQLIAWSVVWTLLI VWGYEFVFQPESLWSRFKKKCFKLTRKMPIIGRK IQDKLNKTKDDISKNMSFLKVDKEYVKALPSQG LSSSAVLEKLKEYSSMDAFWQEGRASGTVYSGE EKLTELLVKAYGDFAWSNPLHPDIFPGLRKIEAEI VRIACSLFNGGPDSCGCVTSGGTESILMACKAYR DLAFEKGIKTPEIVAPQSAHAAFNKAASYFGMKI VRVPLTKMMEVDVRAMRRAISRNTAMLVCSTP QFPHGVIDPVPEVAKLAVKYKIPLHVDACLGGFL IVFMEKAGYPLEHPFDFRVKGVTSISADTHKYGY APKGSSLVLYSDKKYRNYQFFVDTDWQGGIYAS PTIAGSRPGGISAACWAALMHFGENGYVEATKQI IKTARFLKSELENIKGIFVFGNPQLSVIALGSRDFD IYRLSNLMTAKGWNLNQLQFPPSIHFCITLLHAR KRVAIQFLKDIRESVTQIMKNPKAKTTGMGAIYG MAQTTVDRNMGAELSSVFLDSLYSTDTVTQGSQ MNGSPKPH
3092	A	79	2652	LCSQNSPEDWVNFSSEKQKRYPWYWTGRKLRSE RAMKIQKKLTGCSRLMLLCLSLELLLEAGAGNIH YSVPEETDKGSFVGNIAKDLGLQPQELADGGVRI VSRGRMPLFALNPRSGSLITARRIDREELCAQSM PCLVSFNILVEDKMKLFPVEVEIIDINDNTPQFQL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EELEFKMNEITTPGTRVSLPFGQDLDVGMNSLQSYQLSSNPHFSLDVQQGADGPQHPEMVLQSPLDREEEAVHHLILTASDGGEPVRSGTLRIYIQVVDANDNPPAFTQAQYHINVPENVPLGTQLLMVNATDPDEGANGEVTYSFHNVDHRVAQIFRLDSYTGEISNKEPLDFEEYKMYSMEVQAQDGAGLMAKVKVLIKVLDVNDNAPEVTITSVTTAVPENFPPGTIIALISVHDQDSGDNGYTTCFIPGNLPFKLEKLVDNYYRLVTERTLDRELISGYNITITAIDQGTPALSTETHISLLVTDINDNSPVFHQDSYSAYIPENNPRGASIFSVRAHDLDSNENAQITYSLIEDTIQGAPLSAYLSINSDTGVLYALRSFDYEQFRDMQLKVMARDSGDPPLSSNVSLSLFLLDQNDNAPEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVGLHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTLTVAVADRIPDILADLGSLEPSAKPNDSDLTLYLVVAEAAVSCVFLAFVIVLLAHRLRRWHKSRLLQASGGGLASTPGSHFVGVDGVRAFLQTYSHEVSLTADSRKSHLIFPQPNYADTLISQESCEKKGFLSAPQSLLEDKKEPFSQVNFCDECISYLEKNNS
3093	A	1	3868	PPDNQKLGLLEALLKIGDWQHAQNIMDQMPPYYAASHKLIALAICKLIHITIEPLYRSVTSWAVDHAGFLESDPCDSTVGHLLSRVGVPKGAKGSPVNALQNKRAPKQAESFEDLRRDVFNMFCYLGPHLSHDPILFAKVVRIGKSFMKEFQSDGSKQEDKEKTEVILSCLLSITDQVLLPSLSLMDCNACMSEELWGMFKTFPYQHRYRLYGQWKNETYNSHPLLVKVKAQTIDRAKYIMKRLTKENVKPSGRQIGKLSHSNPTILFDYVCFEILSQIQKYDNLITPVVDSLKYLTSLNYDVLACILSNCIIEALANPEKERMKHDDTTISSWLQSLASFCGAVFRKYPIDLAGLLQYVANQLKAGKSFDLLILKEVVQKMAGIEITEEMTMEQLEAMTGGEQLKAEGGYFGQIRNTKKSSQRLKDALLDHDLALPLCLLMAQQRNGVIFQEGGEKHLKLVGKLYDQCHDTLVQFGGFLASNLSTEDYIKRVPSIDVLCNEFHTPHDAAFFLSRPMYAHHISSKYDELKKSEKGSKQQHKVHKYITSCEMVMAPVHEAVVSLHVSKVWDDISPQFYATFWSLTMYDLAVPHTSYEREVNKLKVQMKAIDDNQEMPPNKKKKEKERCTALQDKLLEEEKKQMEHVQRVLQRLKLEKDNWLLAKSTKNETITKFLQLCIFPRCIFSAIDAVYCARFVELVHQQKTPNFSTLLCYDRVFSDIIYTVASCTENEASRYGRFLCCMLETVTRWHSDRATYEKECGNYPGFLTILRATGFDGGNKADQLDYENFRHVVHKWHYKLTKASVHCLETGEYTHIRNILIVLTKILPWYPKVLNLGQALERRVHKICQEEKEKRPDLYALAMGYSGQLKSRKSYMIPENEFHHKDPPPRNAVASVQNGPGGGPSSSSIGSASKSDESSTEETDKSRERSQCGVKAVNKASSTTPKGNSSNGNSGSNSNKAVKENDKEKGKEKEKEKKEKTPATTPEARVLGKDGKEKPKEERPNKDEKARETKERTPKSDKEKEKFKKEEKAKDEKFKTTVPNAESKSTQEREREKEPSRERDIAKEMKSKENVKGGEKTPVSGSLKSPVPRSDIPEPEREQKRRKIDTHPSPSHSSTVKDSLIELKESSAKLYINHTPP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PLSKSKEREMDKKDLDKSRERSREREKKDEKDR KERKRDHSNNDREVPPDLTKRRKEENGTMGVSK HKSESPCESPYPNEKDKEKNKSKSSGKEKGSDSF KSEKMDKISSGGKKESRHDKEKIEKKEKRDSSGG KEEKKHHKSSDKHR
3094	A	2	891	AMLGTREPSRRGAGAVQAEVSERLAMAGPQQQ PPYLHLAELTASQFLEIWKHFDADGNGYIEGKEL ENFFQELEKARKGSGMMSKSDNFGEKMKEFMQ KYDKNSDGKIEMAELAQILPTEENFLLCFRQHVG SSAEFMEAWRKYDTDRSGYIEANELKGFLSDLL KKANRPYDEPKLQEYTQTILRMFDLNGDGKLGL SEMSRLLPVQENFLLKFQGMKLTSEEFNAIFTFY DKDRSGYIDEHELDALLKDLYEKNKKEMNIQQL TNYRKSVMSLAEAGKLYRKDLEIVLCSEPPM
3095	A	1685	700	RRPTGRPGALGAPAAGRVGMPLHVKWPFPAVPP LTWTLASSVVMGLVGTYSCFWTKYMNHLTVHN REVLYELIEKRGPATPLITVSNHQSCMDDPHLWG ILKLRHIWNLKLMRWTPAAADICFTKELHSHFFS LGKCVPVCRGAEFFQAENEGKGVLDTGRHMPG AGKRREKGDGVYQKGMDFILEKLNHGDWVHIF PEGKVNMSSEFLRFKWGIGRLIAECHLNPIILPLW HVGMNDVLPNSPPYFPRFGQKITVLIGKPFSALP VLERLRAENKSAVEMRKALTDFIQEEFQHLKTQ AEQLHNHLQAWEIGLACCLLDSWPAQSWG
3096	A	6642	4022	FVPGLREPQWEPAQPSATMSAPSEEEEYARLVM EAQPEWLRAEVKRLSHELAETTREKIQAAEYGL AVLEEKHQLKLQFEELEVDYEAIRSEMEQLKEAF GQAHTNHKKVAADGESREESLIQESASKEQYYV RKVLELQTELKQLRNVLTNTQSENERLASVAQE LKEINQNVEIQRGRLRDDIKEYKFREARLLQDYS ELEEENISLQKQVSVLRQNQVEFEGLKHEIKRLE EETEYLNSQLEDAIRLKEISERQLEEALETLKTER EQKNSLRKELSHYMSINDSFYTSHLHVSLDGLKF SDDAAEPNNDAEALVNGFEHGGLAKLPLDNKTS TPKKEGLAPPSPSLVSDLLSELNISEIQKLKQQLM QMEREKAGLLATLQDTQKQLEHTRGSLSEQQEK VTRLTENLSALRRLQASKERQTALDNEKDRDSH EDGDYYEVDINGPEILACKYHVAVAEAGELREQ LKALRSTHEAREAQHAEEKGRYEAEGQALTEKV SLLEKASRQDRELLARLEKELKKVSDVAGETQG SLSVAQDELVTFSEELANLYHHVCMCNNETPNR VMLDYYREGQGGAGRTSPGGRTSPEARGRRSPI LLPKGLLAPEAGRADGGTGDSSPSPGSSLPSPLSD PRREPMNIYNLIAIIRDQIKHLQAAVDRTTELSRQ RIASQELGPAVDKDKEALMEEILKLKSLLSTKRE QITTLRTVLKANKQTAEVALANLKSKYENEKAM VTETMMKLRNELKALKEDAATFSSLRAMFATRC DEYITQLDEMQRQLAAAEDEKKTLNSLLRMAIQ QKLALTQRLELLELDHEQTRRGRAKAAPKTKPA TPSVSHTCACASDRAEGTGLANQVFCSEKHSIYC D
3097	A	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKITA SERLRRRPRATARLRAHAAPPEPPLAVFAPPSDR KELLALPVACDPVIASVMSWVQAASLIQGPGDK GDVFDEEADESLLAQREWQSNMQRRVKEGYRD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GIDAGKAVTLQQGFNQGYKKGAEVILNYGRLRGTLSALLSWCHLHNNNSTLINKINNLLDAVGQCEEYVLKHLKSITPPSHVVDLLDSIEDMDLCHVVPAEKKIDEAKDERLCENNAEFNKNCSKSHSGIDCSYVECCRTQEHAHSGKPKPHMDFGTDSQF
3098	A	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYLSADRRVLGLREWGRPASERECSLCQRLKRELNMGDVEKGKKIFIMKCSQCHTVEKGGKHKTGPNLHGLFGRKTGQAPGYSYTAANKNKGIIWGEDTLMEYLENPKKYIPGTKMIFVGIKKKEERADLIAYLKKATNE
3099	A	144	1386	WAVGQARSFPSHPRMSSWIWSRRWSPSVALRVTCTSTSSQRWTVLALSKPGSQQQVSMHTPAPGPPTAGHTEPPSEPPRRARVAKYRAKFDPRVTAKYDIKALIGRGSFSRVVRVEHRATRQPYAIKMIETKYREGREVCESELRVLRRVRHANIIQLVEVFETQERVYMVMELATGGELFDRIIAKGSFTERDATRVLQMVLDGVRYLHALGITHRDLKPENLLYYHPGTDSKIIITDFGLASARKKGDDCLMKTTCGTPEYIAPEVLVRKPYTNSVDMWALGVIAYILLSGTMPFEDDNRTRLYRQILRGKYSYSGEPWPSVSNLAKDFIDRLLTVDPGARMTALQALRHPWVVSMAASSSMKNLHRSISQNLLKRASSRCQSTKSAQSTRSSRSTRSNKSRRVRERELREL
3100	A	3	1500	ARWNGRWVQVPAWPGPGCGTNASGERQRQLPRAWRPVGRTLGSEPIALAWSPPLYLFPIPLPSWAVSQPTPTLGTMFADLDYDIEEDKLGIPTVPGKVTLQKDAQNLIGISIGGGAQYCPCLYIVQVFDNTPAALDGTVAAGDEITGVNGRSIKGKTKVEVAKMIQEVKGEVTIHYNKLQADPKQGMSLDIVLKKVKHRLVENMSSGTADALGLSRAILCNDGLVKRLEELERTAELYKGMTEHTKNLLRAFYELSQTHRGNGIPQSCAFGDVFSVIGVREPQPAASEAFVKFADAHRSIEKFGIRLLKTIKPMLTDLNTYLNKAIPDTRLTIKKYLDVKFEYLSYCLKVKEMDDEEYSCIALGEPLYRVSTGNYEYRLILRCRQEARARFSQMRKDVLEKMELLDQKHVQDIVFQLQRLVSTMSKYYNDCYAVLRDADVFPIEVDLAHTTLAYGLNQEEFTDGEEEEEEDTAAGEPSRDTRGAAGPLDKGGSWCDS
3101	A	1173	197	QGMDSKQQCVKLNDGHFMPVLGFGTYAPPEVPRSKALEVTKLAIEAGFRHIDSAHLYNNEEQVGLAIRSKIADGSVKREDIFYTSKLWSTFHRPELVRPALENSLKKAQLDYVDLYLIHSPMSLKPGEELSPTDENGKVIFDIVDLCTTWEAMEKCKDAGLAKSIGVSNFNRRQLEMILNKPGLKYKPVCNQVECHPYFNRSKLLDFCKSKDIVLVAYSALGSQRDKRWVDPNSPVLLEDPVLCALAKKHKRTPALIALRYQLQRGVVVLAKSYNEQRIRQNVQVFEFQLTAEDMKAIDGLDRNLHYFNSDSFASHPNYPYSDEY
3102	A	144	1098	EQPRPPPCGRRPLPLGSAPCRVRLGRAPRQAPAMSMLPSFGFTQEQVACVCEVLQQGGNLERLGRFLWSLPACDHLHKNESVLKAKAVVAFHRGNFRELYKILESHQFSPHNHPKLQQLWLKAHYVEAEKLRGRPLGAVGKYRVRQKFPLPRTIWDGEETSYCFKEKSRGVLREWYAHNPYPSPREKRELAEATGLTT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TQVSNWFKNRRQRDRAAEAKERENTENNNSSSN KQNQLSPLEGGKPLMSSSEEEFSPPQSPDQNSVLL LQGNMGHARSSNYSLPGLTASQPSHGLQTHQHQ LQDSLLGPLTSSLVDLGS
3103	A	111	1582	LVYSWGCHIMADNDTDRNQTEKLLKRVRELEQ EVQRLKKEQAKNKEDSNIRENSSGAGKTKRAFD FSAHGRRHVALRIAYMGWGYQGFASQENTNNTI EEKLFEALTKTRLVESRQTSNYHRCGRTDKGVS AFGQVISLDLRSQFPRGRDSEDFNVKEEANAAAE EIRYTHILNRVLPPDIRILAWAPVEPSFSARFSCLE RTYRYFFPRADLDIVTMDYAAQKYVGTHDFRNL CKMDVANGVINFQRTILSAQVQLVGQSPGEGRW QEPFQLCQFEVTGQAFLYHQVRCMMAILFLIGQ GMEKPEIIDELLNIEKNPQKPQYSMAVEFPLVLY DCKFENVKWIYDQEAQEFNITHLQQLWANHAV KTHMLYSMLQGLDTVPVPCGIGPKMDGMTEWG NVKPSVIKQTSAFVEGVKMRTYKPLMDRPKCQG LESRIQHFVRRGRIEHPHLFHEEETKAKRDCNDT LEEDNTNLETPTKRVCVDTEIKSII
3104	A	227	1519	VTLIKMNAMLETPELPAVFDGVKLAAVAAVLYV IVRCLNLKSPTAPPDLYFQDSGLSRFLLKSCPLLT KEYIPPLIWGKSGHIQTALYGKMGRVRSPHPYGH RKFITMSDGATSTFDLFEPLAEHCVGDDITMVICP GIANHSEKQYIRTFVDYAQKNGYRCAVLNHLGA LPNIELTSPRMFTYGCTWEFGAMVNYIKKTYPLT QLVVVGFSLGGNIVCKYLGETQANQEKVLCCVS VCQGYSALRAQETFMQWDQCRRFYNFLMADN MKKIILSHRQALFGDHVKKPQSLEDTDLSRLYTA TSLMQIDDNVMRKFHGYNSLKEYYEEESCMRYL HRIYVPLMLVNAADDPLVHESLLTIPKSLSEKRE NVMFVLPLHGGHLGFFEGSVLFPEPLTWMDKLV VEYANAICQWERNKLQCSDTEQVEADLE
3105	A	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPEPP ADEAARAGEGFRYIKPVPGLLLREYLYGGGRDE EPSGAAPEGGATPTAAPETPAPPTRETCYFLNATI LFLFRELRDTALTRRWVTKKIKVEFEELLQTKTA GRLLEGLSLRDVFLGETVPFIKTIRLVRPVVPSAT GEPDGPEGEALPAACPEELAFEAEVEYNGGFHLA IDVDLVFGKSAYLFVKLSRVVGRLRLVFTRVPFT HWFFSFVEDPLIDFEVRSQFEGRPMPQLTSIIVNQ LKKIIKRKHTLPNYKIRFKPFFPYQTLQGFEEDEE HIHIQQWALTEGRLKVTLLECSRLLIFGSYDREA NVHCTLELSSSVWEEKQRSSIKTGTISLTAVFMG WHRVSEAFPGLWYKLLVDLPFWGLEDGGPLLT VPLRQCPG
3106	A	972	468	MAAAGAGRLRRVASALLLRSPRLPARELSAPAR LYHKKVVDHYENPRNVGSLDKTSKNVGTGLVG APACGDVMKLQIQVDEKGKIVDARFKTFGCGSA IASSSLATEWVKGKTVEEALTIKNTDIAKELCLPP VKLHCSMLAEDAIKAALADYKLKQEPKKGEAE KK
3107	A	106	1221	TCQDVRSVFSLVRANIFGEESTAGAGWHREEDM RKELQLSLSVTLLLVCGFLYQFTLKSSCLFCLPSF KSHQGLEALLSHRRGIVFLETSERMEPPHLVSCS VESAAKIYPEWPVVFFMKGLTDSTPMPSNSTYPA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				FSFLSAIDNVFLFPLDMKRLLEDTPLFSWYNQINASAERNWLHISSDASRLAIIWKYGGIYMDTDVISIRPIPEENFLAAQASRYSSNGIFGFLPHHPFLWECMENFVEHYNSAIWGNQGPELMTRMLRVWCKLEDFQEVSDLRCLNISFLHPQRFYPISYREWRRYYEVWDTEPSFNVSYALHLWNHMNQEGRAVIRGSNTLVENLYRKHCPRTYRDLIKGPEGSVTGELGPGNK
3108	A	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQRNFQLMRDLDQRTEDLKAEIDKLATEYMSSARSLSSEEKLALLKQIQEAYGKCKEFGDDKVQLAMQTYEMVDKHIRRLDTDLARFEADLKEKQIESSDYDSSSSKGKKKGRTQKEKKAARARSKGKNSDEEAPKTAQKKLKLVRTSPEYGMPSVTFGSVHPSDVLDMPVDPNEPTYCLCHQVSYGEMIGCDNPDCSIEWFHFACVGLTTKPRGKWFCPRCSQERKKK
3109	A	1	2613	MVAVRAAGPREGASQDEAGTVWAPMTGCPCQCRPGPSWLLVDTLEPETAYPVQRPGPEQAGNQRLQMKRAQFGPHDWLSLPVPPGPSWLLVDTLEPETAYQFSVLAQNKLGTSAFSEVVTVNTLAFPITTPEPLVLVTPPRCLIANRTQQGVLLSWLPPANHSFPIDRYIMEFRVAERWELLDDGIPGTEGEFFAKDLSQDTWYEFRVLAVMQDLISEPSNIAGVSSTDIFPQPDLTEDGLARPVLAGIVATICFLAAAILFSTLAACFVNKQRKRKLKRKKDPPLSITHCRKSLESPLSSGKVSPESIRTLRAPSESSDDQGQPAAKRMLSPTREKELSLYKKTKRAISSKKYSVAKAEAEAEATTPIELISRGPDGRFVMDPAEMEPSLKSRRIEGFPFAEETDMYPEFRQSDEENEDPLVPTSVAALKSQLTPLSSSQESYLPPPAYSPRFQPRGLEGPGGLEGRLQATGQARPPAPRPFHHGQYYGYLSSSSPGEVEPPPFYVPEVGSPLSSVMSSPPLPTEGPFGHPTIPEENGENASNSTLPLTQTPTGGRSPEPWGRPEFPFGGLETPAMMFPHQLPPCDVPESLQPKAGLPRGLPPTSLQVPAAYPGILSLEAPKGWAGKSPGRGPVPAPPAAKWQDRPMQPLVSQGQLRHTSQGMGIPVLPYPEPAEPGAHGGPSTFGLDTRWYEPQPRPRPSPRQARRAEPSLHQVVLQPSRLSPLTQSPLSSRTGSPELAARARPRPGLLQQAEMSEITLQPPAAVSFSRKSTPSTGSPSQSSRSGSPSYRPAMGFTTLATGYPSPPPGPAPAGPGDSLDVFGQTPSPRRTGEELLRPETPPPTLPTLGKLRRDRPAPATSPPERALSKL
3110	A	88	924	ILGSRTMSLTNTKTGFSVKDILDLPDTNDEEGSVAEGPEEENEGPEPAKRAGPLGQGALDAVQSLPLKNPFYDSSDNPYTRWLASTEGLQYSLHGLAAGAPPQDSSSKSPEPSADESPDNDKETPGGGGDAGKKRKRRVLFSKAQTYELERRFRQQRYLSAPEREHLASLIRLTPTQVKIWFQNHRYKMKRARAEKGMEVTPLPSPRRVAVPVLVRDGKPCHALKAQDLAAATFQAGIPFSAYSAQSLQHMQYNAQYSSASTPQYPTAHPLVQAQQWTW
3111	A	595	291	PSVASLARRFSGRALWPPSHSVPGNRALCPRLLHGTTLPGGNQRELARQKNMKKQSDSVKGKRRDDGLSAAARKQRDSTPRDSEIMQQKQKKANEKKEEPK
3112	A	3641	1555	APMLQIHHFSFKLIFQNIHKSKFISQRLSQNADST

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RHTNLSNTHYSDLIVWNCCLFFRNWCNEFFLKS CHFAQEREGSGDLCNSRAEKTKSAACVIFRRFPV APLIPYPLITKEDINAIEMEEDKRDLISREISKFRDT HKKLEEEKGKKEKERQEIEKERRERERERERERE RRERERERERERERERKEKERERERERDRDRDRTK ERDRDRDRERDRDRDRERSSDRNKDRSRSREKS RDRERERERERERERERERERERERERERERERE REREKDKKRDREEDEEDAYERRKLERKLREKEA AYQERLKNWEIRERKKTREYEKEAEREEERRRE MAKEAKRLKEFLEDYDDDRDDPKYYRGSALQK RLRDREKEMEADERDRKREKEELEEIRQRLLAE GHPDPDAELQRMEQEAERRRQPQIKQEPESEEEE EEKQEKEEKREEPMEEEEEPEQKPCLKPTLRPISS APSVSSASGNATPNTPGDESPCGIIIPHENSPDQQ QPEEHRPKIGLSLKLGASNSPGQPNSVKRKKLPV DSVFNKFEDEDSDDVPRKRKLVPLDYGEDDKNA TKGTVNTEEKRKHIKSLIEKIPTAKPELFAYPLDW SIVDSILMERRIRPWINKKIIEYIGEEEATLVDLVC SKVMAHSPPQSILDDVAMVLDEEAEVFIVKMWR LLIYETEAKKIGLVK
3113	A	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPGLET NILKMTTPNKTPPGADPKQLERTGTVREIGSQAV WSLSSCKPGFGVDQLRDDNLETYWQSDGSQPHL VNIQFRRKTTVKTLCIYADYKSDESYTPSKISVRV GNNFHNLQEIRQLELVEPSGWIHVPLTDNHKKPT RTFMIQIAVLANHQNGRDTHMRQIKIYTPVEESSI GKFPRCTTIDFMMYRSIR
3114	A	1	1613	MTSKEESRRQQPTAGPAGQGKLPSPSEPQLPTPP TRSLHHFRRPLSPSREAQAHIAPSSELHLPQSQSA GPPPLGAGTEVELVVPGRDEGSRGALPGSSGVKF VWRKIVRFPVSDQVRTLSISRLMRRLLEMMQTL VQFIIGWRSLLGRTLGTIMNTMYVMMAQILRSH LIKATVIPNRVKMLPYFGIIRNRMMSTHKSKKKI REYYRLLNVEEGCSADEVRESFHKLAKQYHPDS GSNTADSATFIRIEKAYRKVLSHVIEQTNASQSK GEEEEDVEKFKYKTPQHRHYLSFEGIGFGTPTQR EKHYRQFRADRAAEQVMEYQKQKLQSQYFPDS VIVKNIRQSKQQKITQAIERLVEDLIQESMAKGDF DNLSGKGKPLKKFSDCSYIDPMTHNLNRILIDNG YQPEWILKQKEISDTIEQLREAILVSRKKLGNPMT PTEKKQWNHVCEQFQENIRKLNKRINDFNLIVPI LTRQKVHFDAQKEIVRAQKIYETLIKTKEVTDRN PNNLDQGEGEKTPEIKKGFLNLMDLVEIY
3115	A	1	2036	FRHRCGCLSYCRSRRGIRRVEPLRRARARVGPRF RPLCRMEIIRSNFKSNLHKVYQAIEEADFFAIDGE FSGISDGPSVSALTNGFDTPEERYQKLKKHSMDF LLFQFGLCTFKYDYTDSKYITKSFNFYVFPKPFNR SSPDVKFVCQSSSIDFLASQGFDFNKGFRKGIPYL NQEEERQLREQYDEKRSQANGAGALSYVSPNTS KCPVTIPEDQKKFIDQVVEKIEDLLQSEENKNLDL EPCTGFQRKLIYQTLSWKYPKGIHVETLETEKKE RYIVISKVDEEERKRREQQKHAKEQEELNDAVG FSRVIHAIANSGKLVIGHNMLLDVMHTVHQFYC PLPADLSEFKEMTTCVFPRLLDTKLMASTQPFKD IINNTSLAELEKRLKETPFNPPKVESAEGFPSYDT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ASEQLHEAGYDAYITGLCFISMANYLGSFLSPPKIHVSARSKLIEPFFNKLFLMRVMDIPYLNLEGPDLQPKRDHVLHVTFPKEWKTSDLYQLFSAFGNIQISWIDDTSAFVSLSQPEQVKIAVNTSKYAESYRIQTYAEYMGRKQEEKQIKRKWTEDSWKEADSKRLNPQCIPYTLQNHYYRNNSFTAPSTVGKRNLSPSQEEAGLEDGVSGEISDTELEQTDSCAEPLSEGRKKAKKLKRMKKELSPAGSISKNSPATLFEVPDTW
3116	A	3	1443	TREAPMALAVAPWGRQWEEARALGRAVRMLQRLEEQCVDPRLSVSPPSLRDLLPRTAQLLREVAHSRRAAGGGGPGGPGGSGDFLLIYLANLEAKSRQVAALLPPRGRRSANDELFRAGSRLRRQLAKLAIIFSHMHAELHALFPGGKYCGHMYQLTKAPAHTFWRESCGARCVLPWAEFESLLGTCHPVEPGCTALALRTTIDLTCSGHVSIFEFDVFTRLFQPWPTLLKNWQLLAVNHPGYMAFLTYDEVQERLQACRDKPGSYIFRPSCTRLGQWAIGYVSSDGSILQTIPANKPLSQVLLEGQKDGFYLYPDGKTHNPDLTELGQAEPQQRIHVSEEQLQLYWAMDSTFELCKICAESNKDVKIEPCGHLLCSCCLAAWQHSDSQTCPFCRCEIKGWEAVSIYQFHGQATAEDSGNSSDQEGRELELGQVPLSAPPLPPRPDLPPRKPRNAQPKVRLLKGNSPPAALGPQDPAPA
3117	A	296	3547	ERHSSPLLQHILTHALMRNKKHSNNWLAQHWFQSSIILCFSPVGRTLRVRARKFPAIVNCTAIDWFHAWPQEALVSVSRRFIEETKGIEPVHKDSISLFMAHVHTTVNEMSTRYYQNERRHNYTTPKSFLEQISLFKNLLKKKQNEVSEKKERLVNGIQKLKTTASQVGDLKARLASQEAELQLRNHDAEALITKIGLQTEKVSREKTIADAEERKVTAIQTEVFQKQRECEADLLKAEPALVAATAALNTLNRVNLSELKAFPNPPIAVTNVTAAVMVLLAPRGRVPKDRSWKAAKVFMGKVDDFLQALINYDKEHIPENCLKVVNEHYLKDPEFNPNLIRTKSFAAAGLCAWVINIIKFYEVYCDVEPKRQALAQANLELAAATEKLEAIRKKLVVSANYDIEKSEKIRWGQSIKSFEAQEKTLCGDVLLTAAFVSYVGPFTRQYRQELVHCKWVPFLQQKVSIPLTEGLDLISMLTDDATIAAWNNEGLPSDRMSTENAAILTHCERWPLVIDPQQQGIKWIKNKYGMDLKVTHLGQKGFLNAIETALAFGDVILIENLEETIDPVLDPLLGRNTIKKGKYIRIGDKECEFNKNFRLILHTKLANPHYKPELQAQTTLLNFTVTEDGLEAQLLAEVVSIERPDLEKLKLVLTKHQNDFKIELKYLEDDLLLRLSAAEGSFLDDTKLVERLEATKTTVAEIEHKVIEAKENERKINEARECYRPVAARASLLYFVINDLQKINPLYQFSLKAFNVLFHRAIEQADKVEDMQGRISILMESITHAVFLYTSQALFEKDKLTFLSQMAFQILLRKKEIDPLELDFLLRFTVEHTHLSPVDFLTSQSWSAIKAIAVMEEFRGIDRDVEGSAKQWRKWVESECPEKEKLPQEWKKKSLIQKLILLRAMRPDRMTYALRNFVEEKLGAKYVERTRLDLVKAFEESSPATPIFFILSPGVDALKDLEILGKRLGFTIDSGKFHNVSLGQGQETVAEVALEKASKGGHWVILQNVHLVAKWLGTLEKLLERFSQGSHRDYRVFMSAESAPTPDEHIIPQGLLENSIKITNEPPTGMLANLHAALYNFD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				Q
3118	A	1	226	PYSLSTSCLGSPTSPRLEMDPNCSCATGGSCTCTGSCKCKECKCNSCKKSECGAISRNLGLSQVRGRKPELGMEE
3119	A	1254	4133	PLATLTMEEQGHSEMEIIPSESHPHIQLLKSNRELLVTHIRNTQCLVDNLLKNDYFSAEDAEIVCACPTQPDKVRKILDLVQSKGEEVSEFFLYLLQQLADAYVDLRPWLLEIGFSPSLLTQSKVVVNTDPVSRYTQQLRHHLGRDSKFVLCYAQKEELLLEEIYMDTIMELVGFSNESLGSLNSLACLLDHTTGILNEQGETIFILGDAGVGKSMLLQRLQSLWATGRLDAGVKFFFHFRCRMFSCFKESDRLCLQDLLFKHYCYPERDPEEVFAFLLRFPHVALFTFDGLDELHSDLDLSRVPDSSCPWEPAHPLVLLANLLSGKLLKGASKLLTARTGIEVPRQFLRKKVLLRGFSPSHLRAYARRMFPERALQDRLLSQLEANPNLCSLCSVPLFCWIIFRCFQHFRAAFEGSPQLPDCTMTLTDVFLLVTEVHLNRMQPSSLVQRNTRSPVETLHAGRDTLCSLGQVAHRGMEKSLFVFTQEEVQASGLQERDMQLGFLRALPELGPGGDQQSYEFFHLTLQAFFTAFFLVLDDRVGTQELLRFFQEWMPPAGAATTSCYPPFLPFQCLQGSGPAREDLFKNKDHFQFTNLFLCGLLSKAKQKLLRHLVPAAALRRKRKALWAHLFSSLRGYLNSLPRVQVESFNQVQAMPTFIWMLRCIYETQSQKVGQLAARGICANYLKLTYCNACSADCSALSFVLHHFPKRLALDLDNNNLNDYGVRELQPCFSRLTVLRLSVNQITDGGVKVLSEELTKYKIVTYLGLYNNQITDVGARYVTKILDECKGLTHLKLGKNKITSEGGKYLALAVKNSKSISEVGMWGNQVGDEGAKAFAEALRNHPSLTTLSLASNGISTEGGKSLARALQQNTSLEILWLTQNELNDEVAESLAEMLKVNQTLKHLWLIQNQITAKGTAQLADALQSNTGITEICLNGNLIKPEEAKVYEDEKRIICF
3120	A	43	1004	QLWGFAAGSDSRPAMGCDGGTIPKRHELVKGPKKVEKVDKDAELVAQWNYCTLSQEILRRPIVACELGRLYNKDAVIEFLLDKSAEKALGKAASHIKSIKNVTELKLSDNPAWEGDKGNTKGDKHDDLQRARFICPVVGLEMNGRHRFCFLRCCGCVFSERALKEIKAEVCHTCGAAFQEDDVIVLNGTKEDVDVLKTRMEERRLRAKLEKKTKKPKAAESVSKPDVSEEAPGPSKVKTGKPEEASLDSREKKTNLAPKSTAMNESSSGKAGKPPCGATKRSIADSEESEAYKSLFTTHSSAKRSKEESAHWVTHTSYCF
3121	A	3	1490	HASGPTRPVSWSFHKLKTMKHLLLLLLCVFLVKSQGVNDNEEGFFSARGHRPLDKKREEAPSLRPAPPPISGGGYRARPAKAAATQKKVERKAPDAGGCLHADPDLGVLCPTGCQLQEALLQQERPIRNSVDELNNNVEAVSQTSSSSFQYMYLLKDLWQKRQKQVKDNENVVNEYSSELEKHQLYIDETVNSNIPTNLRVLRSILENLRSKIQKLESDVSAQMEYCRTPCTVSCNIPVVSGKECEEIIRKGGETSEMYLIQPDSSVKPYRVYCDMNTENGGWTVIQNRQDGSVDFGRKWDPYKQGFGNVATNTDGKNYCGLPGEYWLGNDKISQLTRMGPTELLIEMEDWKGDKVKAHYGGFTVQNEANKYQISVNKYRGTAGNALMDGASQLMGE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NRTMTIHNGMFFSTYDRDNDGWLTSDPRKQCSK EDGGGWWYNRCHAANPNGRYYWGGQYTWDM AKHGTDDGVVWMNWKGSWYSMKKMSMKIRP FFPQQ
3122	A	3	1490	HASGPTRPVSWSFHKLKTMKHLLLLLLCVFLVK SQGVNDNEEGFFSARGHRPLDKKREEAPSLRPAP PPISGGGYRARPAKAAATQKKVERKAPDAGGCL HADPDLGVLCPTGCQLQEALLQQERPIRNSVDEL NNNVEAVSQTSSSSFQYMYLLKDLWQKRQKQV KDNENVVNEYSSELEKHQLYIDETVNSNIPTNLR VLRSILENLRSKIQKLESDVSAQMEYCRTPCTVS CNIPVVSGKECEEIIRKGGETSEMYLIQPDSSVKP YRVYCDMNTENGGWTVIQNRQDGSVDFGRKW DPYKQGFGNVATNTDGKNYCGLPGEYWLGNDK ISQLTRMGPTELLIEMEDWKGDKVKAHYGGFTV QNEANKYQISVNKYRGTAGNALMDGASQLMGE NRTMTIHNGMFFSTYDRDNDGWLTSDPRKQCSK EDGGGWWYNRCHAANPNGRYYWGGQYTWDM AKHGTDDGVVWMNWKGSWYSMKKMSMKIRP FFPQQ
3123	A	3	1490	HASGPTRPVSWSFHKLKTMKHLLLLLLCVFLVK SQGVNDNEEGFFSARGHRPLDKKREEAPSLRPAP PPISGGGYRARPAKAAATQKKVERKAPDAGGCL HADPDLGVLCPTGCQLQEALLQQERPIRNSVDEL NNNVEAVSQTSSSSFQYMYLLKDLWQKRQKQV KDNENVVNEYSSELEKHQLYIDETVNSNIPTNLR VLRSILENLRSKIQKLESDVSAQMEYCRTPCTVS CNIPVVSGKECEEIIRKGGETSEMYLIQPDSSVKP YRVYCDMNTENGGWTVIQNRQDGSVDFGRKW DPYKQGFGNVATNTDGKNYCGLPGEYWLGNDK ISQLTRMGPTELLIEMEDWKGDKVKAHYGGFTV QNEANKYQISVNKYRGTAGNALMDGASQLMGE NRTMTIHNGMFFSTYDRDNDGWLTSDPRKQCSK EDGGGWWYNRCHAANPNGRYYWGGQYTWDM AKHGTDDGVVWMNWKGSWYSMKKMSMKIRP FFPQQ
3124	A	3	544	RVDDFVLLRSRLALRWLSHVRRPSRRVPRMPRG SRSRTSRMAPPASRAPQMRAAPRPAPVAQPPAA APPSAVGSSAAAPRQPGLMAQMATTAAGVAVG SAVGHTLGHAITGGFSGGSNAEPARPDITYQEPQ GTQPAQQQQPCLYEIKQFLECAQNQGDIKLCEGF NEVLKQCRLANGLA
3125	A	3	571	GNSYNHRSLAAYPYMSHSQHSPYLQSYHNSSAA AQTRGDDTDQQKTTVIENGEIRFNGKGKKIRKPR TIYSSLQLQALNHRFQQTQYLALPERAELAASLG LTQTQVKIWFQNKRSKFKKLLKQGSNPHESDPL QGSAALSPRSPALPPVWDVSASAKGVSMPPNSY MPGYSHWYSSPHQDTMQRPQMM
3126	A	43	5377	LSVFFPIPVDGRDRGSNPSLESTSSELSTSTSEGSL SAMSGRNELHSRLHPHPQSSLIPMMFSPPESLLAS CILRGNFAEAHQVLFTFNLKSSPSSGELMFMERY QEVIQELAQVEHKIENQNSDAGSSTIRRTGSGRST LQAIGSAAAAGMVFYSISDVTDKLLNTSGDPIPM LQEDFWISTALVEPTAPLREVLEDLSPPAMAAFD LACSQCQLWKTCKQLLETAERRLNSSLERRGRRI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DHVLLNADGIRGFPVVLQQISKSLNYLLMSASQTKSESVEEKGGGPPRCSITELLQMCWPSLSEDCVASHTTLSQQLDQVLQSLREALELPEPRTPPLSSLVEQAAQKAPEAEAHPVQIQTQLLQKNLGKQTPSGSRQMDYLGTFFSYCSTLAAVLLQSLSSEPDHVEVKVGNPFVLLQQSSSQLVSHLLFERQVPPERLAALLAQENLSLSVPQVIVSCCCEPLALCSSRQSQQTSSLLTRLGTLAQLHASHCLDDLPLSTPSSPRTTENPTLERKPYSSPRDSSLPALTSSALAFLKSRSKLLATVACLGASPRLKVSKPSLSWKELRGRREVPLAAEQVARECERLLEQFPLFEAFLLAAWEPLRGSLQQGQSLAVNLCGWASLSTVLLGLHSPIALDVLSEAFEESLVARDWSRALQLTEVYGRDVDDLSSIKDAVLSCAVACDKEGWQYLFPVKDASLRSRLALQFVDRWPLESCLEILAYCISDTAVQEGLKCELQRKLAELQVYQKILGLQSPPVWCDWQTLRSCCVEDPSTVMNMILEAQEYELCEEWGCLYPIPREHLISLHQKHLLHLLERRDHDKALQLLRRIPDPTMCLEVTEQSLDQHTSLATSHFLANYLTTHFYGQLTAVRHREIQALYVGSKILLTLPEQHRASYSHLSSNPLFMLEQLLMNMKVDWATVAVQTLQQLLVGQEIGFTMDEVDSLLSRYAEKALDFPYPQREKRSDSVIHLQEIVHQAADPETLPRSPSAEFSPAAPPGISSIHSPSLRERSFPPTQPSQEFVPPATPPARHQWVPDETESICMVCCREHFTMFNRRHHCRRCGRLVCSSCSTKKMVVEGCRENPARVCDQCYSYCNKDVPEEPSEKPEALDSSKSESPPYSFVVRVPKADEVEWILDLKEEENELVRSEFYYEQAPSASLCIAILNLHRDSIACGHQLIEHCCRLSKGLTNPEVDAGLLTDIMKQLLFSAKMMFVKAGQSQDLALCDSYISKVDVLNILVAAAYRHVPSLDQILQPAAVTRLRNQLLEAEYYQLGVEVSTKTGLDTTGAWHAWGMACLKAGNLTAAREKFSRCLKPPFDLNQLNHGSRLVQDVVEYLESTVRPFVSLQDDDYFATLRELEATLRTQSLSLAVIPEGKIMNNTYYQECLFYLHNYSTNLAIISFYVRHSCLREALLHLLNKESPPEVFIEGIFQPSYKSGKLHTLENLLESIDPTLESWGKYLIAACQHLQKKNYYHILYELQQFMKDQVRAAMTCIRFFSHKAKSYTELGEKLSWLLKAKDHLKIYLQETSRSSGRKKTTFFRKKMTAADVSRHMNTLQLQMEVTRFLHRCESAGTSQITTLPLPTLFGNNHMKMDVACKVMLGGKNVEDGFGIAFRVLQDFQLDAAMTYCRAARQLVEKEKYSEIQQLLKCVSESGMAAKSDGDTILLNCLEAFKRIPPQCCFCSAQELEGLIQAIHNDDNKVRAYLICCKLRSAYLIAVKQEHSRATALVQQVQQAAKSSGDAVVQDICAQWLLTSHPRGAHGPGSRK
3127	A	467	1259	HLGPPLAWIPAASLTSTKGEFGVEDDRPARGPPPPKSEEASWSESGVSSSSGDGPFAGGEVDKRLHQLKTQLATLTSSLATVTQEKSRMEASYLADKKKMKQDLEDASNKAEEERARLEGELKGLQEQIAETKARLITQQHDRAQEQSDHALMLRELQKLLQEERTQRQDLELRLEETREALAGRAYAAEQMEGFELQTKQLTREVEELKSELQAIRDEKNQPDPRLQELQEEAARLKSHFQAQLQQEMRKVIIHISFKHQPLT
3128	A	1854	798	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLHLFL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LTAGPALGWNDPDRMLLRDVKALTLHYDRYTT SRRLDPIPQLKCVGGTAGCDSYTPKVIQCQNKG WDGYDVQWECKTDLDIAYKFGKTVVSCEGYES SEDQYVLRGSCGLEYNLDYTELGLQKLKESGKQ HGFASFSDYYYKWSSADSCNMSGLITIVVLLGIA FVVYKLFLSDGQYSPPPYSEYPPFSHRYQRFTNS AGPPPPGFKSEFTGPQNTGHGATSGFGSAFTGQQ GYENSGPGFWTGLGTGGILGYLFGSNRAATPFSD SWYYPSYPPSYPGTWNRAYSPLHGGSGSYSVCS NSDTKTRTASGYGGTRRR
3129	A	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLVEK CESSVSLTVPPVVKLENGSSTNVSLTLRPPLNATL VITFEITFRSKNITILELPDEVVVPPGVTNSSFQVT SQNVGQLTVYLHGNHSNQTGPRIRFLVIRSSAISII NQVIGWIYFVAWSISFYPQVIMNWRRKSVIGLSF DFVALNLTGFVAYSVFNIGLLWVPYIKEQFLLKY PNGVNPVNSNDVFFSLHAVVLTLIIIVQCCLYERG GQRVSWPAIGFLVLAWLFAFVTMIVAAVGVITW LQFLFCFSYIKLAVTLVKYFPQAYMNFYYKSTEG WSIGNVLLDFTGGSFSLLQMFLQSYNNDQWTLIF GDPTKFGLGVFSIVFDVVFFIQHFCLYRKRPGYD QLN
3130	A	31	2026	CWWPPLLPQLEPEPPPLRPRVAASQGGGMLGKG VVGGGGGTKAPKPSFVSYVRPEEIHTNEKEVTEK EVTLHLLPGEQLLCEASTVLKYVQEDSCQHGVY GRLVCTDFKIAFLGDDESALDNDETQFKNKVIGE NDITLHCVDQIYGVFDEKKKTLFGQLKKYPEKLII HCKDLRVFQFCLRYTKEEEVKRIVSGIIHHTQAP KLLKRLFLFSYATAAQNNTVTDPKNHTVMFDTL KDWCWELERTKGNMKYKAVSVNEGYKVCERL PAYFVVPTPLPEENVQRFQGHGIPIWCWSCHNGS ALLKMSALPKEQDDGILQIQKSFLDGIYKTIHRPP YEIVKTEDLSSNFLSLQEIQTAYSKFKQLFLIDNST EFWDTDIKWFSLLESSSWLDIIRRCLKKAIEITEC MEAQNMNVLLLEENASDLCCLISSLVQLMMDPH CRTRIGFQSLIQKEWVMGGHCFLDRCNHLRQND KEEHQRQLSLPLTQSKSSPKRGFFREETDHLIKNL LGKRISKLINSSDELQDNFREFYDSWHSKSTDYH GLLLPHIEGPEIKVWAQRYLRWIPEAQILGGGQV ATLSKLLEMMEEVQSLQEKIDERHHSQQAPQAE APCLLRNSARLSSLFPFALLQRHSSKPVLPTSGW KALGDEDDLAKREDEFVDLGDV
3131	A	126	965	QSRSRPRREGVGTGSRAVLCILATCGSKMSDIGD WFRSIPAITRYWFAATVAVPLVGKLGLISPAYLF LWPEAFLYRFQIWRPITATFYFPVGPGTGFLYLV NLYFLYQYSTRLETGAFDGRPADYLFMLLFNWI CIVITGLAMDMQLLMIPLIMSVLYVWAQLNRDM IVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIG NLVGHLYFFLMFRYPMDLGGRNFLSTPQFLYRW LPSRRGGVSGFGVPPASMRRAADQNGGGGRHN WGQGFRLGDQ
3132	A	2	350	FVAGWRALTAPSTSARLRAFGWQAAARLLVFG ARGVGLGSGAPGSLPCYLRMDALALLGGLVNV ARLPERWGPGRFDYWGNSHQIMHLLSVGSILQL HAGVVPDLLWAAHHACPRD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3133	A	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPNLYSQLNALQFTVDERSILWLNQFLLDLKQSLNQFMAVYKLNDNSKSDEHVDVRVDGLMLKFVIPSEVKSECHQDQPRAISIQSSEMIATNTRHCPNCRHSDLEALFQDFKDCDFFSKTYTSFPKSCDNFNLLHPIFQRHAHEQDTKMHEIYKGNITPQLNKNTLKTSAATDVWAVYFSQFWIDYEGMKSGKGRPISFVDSFPLSIWICQPTRYAESQKEPQTCNQVSLNTSQSESSDLAGRLKRKKLLKEYYSTESEPLTNGGQKPSSSDTFFRFSPSSSEADIHLLVHVHKHVSMQINHYQYLLLLFLHESLILLSENLRKDVEAVTGSPASQTSICIGILLRSAELALLLHPVDQANTLKSPVSESVSPVVPDYLPTENGDFLSSKRKQISRDINRIRSVTVNHMSDNRSMSVDLSHIPLKDPLLFKSASDTNLQKGISFMDYLSDKHLGKISEDESSGLVYKSGSGEIGSETSDKKDSFYTDSSSVLNYREDSNILSFDSDGNQNILSSTLTSKGNETIESIFKAEDLLPEAASLSENLDISKEETPPVRTLKSQSSLSGKPKERCPPNLAPLCVSYKNMKRSSSQMSLDTISLDSMILEEQLLESDGSDSHMFLEKGNKKNSTTNYRGTAESVNAGANLQNYGETSPDAISTNSEGAQENHDDLMSVVVFKITGVNGEIDIRGEDTEICLQVNQVTPDQLGNISLRHYLCNRPVGSDQKAVIHSKSSPEISLRFESGPGAVIHSLLAEKNGFLQCHIENFSTEFLTSSLMNIQHFLEDETVATVMPMKIQVSNTKINLKDDSPRSSTVSLEPAPVTVHIDHLVVERSDDGSFHIRDSHMLNTGNDLKENVKSDSVLLTSGKYDLKKQRSVTQATQTSPGVPWPSQSANFPEFSFDFTREQLMEENESLKQELAKAKMALAEAHLEKDALLHHIKKMTVE
3134	A	9	1579	EEEGLSGGGPRVPCSLWGKQTMDYDFKAKLAAERERVEDLFEYEGCKVGRGTYGHVYKARRKDGKDEKEYALKQIEGTGISMSACREIALLRELKHPNVIALQKVFLSHSDRKVWLLFDYAEHDLWHIIKFHRASKANKKPMQLPRSMVKSLLYQILDGIHYLHANWVLHRDLKPANILVMGEGPERGRVKIADMGFARLFNSPLKPLADLDPVVVTFWYRAPELLLGARHYTKAIDIWAIGCIFAELLTSEPIFHCRQEDIKTSNPFHHDQLDRIFSVMGFPADKDWEDIRKMPEYPTLQKDFRRTTYANSSLIKYMEKHKVKPDSKVFLLLQKLLTMDPTKRITSEQALQDPYFQEDPLPTLDVFAGCQIPYPKREFLNEDDPEEKGDKNQQQQQNQHQQPTAPPQQAAAPPQAPPPQQNSTQTNGTAGGAGAGVGGTGAGLQHSQDSSLNQVPPNKKPRLGPSGANSGGPVMPSDYQHSSSRLNYQSSVQGSSQSQSTLGYSSSSQQSSQYHPSHQAHRY
3135	A	3	1111	ERKMAEPPSPVHCVAAAAPTATVSEKEPFGKLQLSSRDPPGSLSAKKVRTEEKKAPRRVNGEGGSGGNSRQLQPPAAPSPQSYGSPASWSFAPLSAAPSPSSSRSSFSFSAGTAVPSSASASLSQPGPRKLLVPPTLLHAQPHHLLLPAAAAAASANAKSRRPKEKREKERRRHGLGGAREAGGASREENGEVKPLPRDKIKDKIKERDKEKEREKKKHKVMNEIKKENGEVKILLKSGKEKPKTNIEDLQIKKVKKKKKKKHKENEKRKRPKMYSKSIQTICSGLLTDVEDQAAKGILNDNIKDYVGKNLDTKNYDSKIPENSEFPFVSLKEPRVQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NNLKRLDTLEFKQLIHIEHQPNGGASVIHCLQ
3136	A	1442	682	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEIAC YKNKVVGWRSGVEKDLDEVLQTHSVFVNVSKG QVAKKEDLISAFGTDDQTEICKQILTKGEVQVSD KERHTQLEQMFRDIATIVADKCVNPETKRPYTVI LIERAMKDIHYSVKTNKSTKQQALEVIKQLKEK MKIERAHMRLRFILPVNEGKKLKEKLKPLIKVIES EDYGQQLEIVCLIDPGCFREIDELIKKETKGKGSL EVLNLKDVEEGDEKFE
3137	A	1	3143	MVEGKRHVLHGGRQERMRAKQKGKPLIKSSDL VRLIHYHHNSSPLHKQSSGPSSSPAAAAAPEKPG PKAAEVGDDFLGDFVVGERVWVNGVKPGVVQY LGETQFAPGQWAGVVLDDPVGKNDGAVGGVR YFECPALQGIFTRPSKLTRQPTAEGSGSDAHSVES LTAQNLSLHSGTATPPLTSRVIPLRESVLNSSVKT GNESGSNLSDSGSVKRGEKDLRLGDRVLVGGTK TGVVRYVGETDFAKGEWCGVELDEPLGKNDGA VAGTRYFQCPPKFGLFAPIHKVIRIGFPSTSPAKA KKTKRMAMGVSALTHSPSSSSISSVSSVASSVGG RPSRSGLLTETSSRYARKISGTTALQEALKEKQQ HIEQLLAERDLERAEVAKATSHICEVEKEIALLK AQHEQYVAEAEEKLQRARLLVESVRKEKVDLSN QLEEERRKVEDLQFRVEEESITKGDLETQTQLEH ARIGELEQSLLLEKAQAERLLRELADNRLTTVAE KSRVLQLEEELTLRRGEIEELQQCLLHSGPPPPDH PDAAEILRLRERLLSASKEHQRESGVLRDKYEKA LKAYQAEVDKLRAANEKYAQEVAGLKDKVQQ ATSENMGLMDNWKSKLDSLASDHQKSLEDLKA TLNSGPGAQQKEIGELKAVMEGIKMEHQLELGN LQAKHDLETAMHVKEKEALREKLQEAQEELAG LQRHWRAQLEVQASQHRLELQEAQDQRRDAEL RVHELEKLDVEYRGQAQAIEFLKEQISLAEKKML DYERLQRAEAQGKQEVESLREKLLVAENRLQAV EALCSSQHTHMIESNDISEETIRTKETVEGLQDKL NKRDKEVTALTSQTEMLRAQVSALESKCKSGEK KVDALLKEKRRLEAELETVSRKTHDASGQLVLIS QELLRKERSLNELRVLLLEANRHSPGPERDLSRE VHKAEWRIKEQKLKDDIRGLREKLTGLDKEKSL SDQRRYSLIDPSSAPELLRLQHQLMSTEDALRDA LDQAQQVEKLMEAMRSCPDKAQTIGNSGSANGI HQQDKAQKQEDKH
3138	A	110	2499	QDRRLLRLELQKTCQPTSTMSGSHTPACGPFSAL TPSIWPQEILAKYTQKEESAEQPEFYYDEFGFRV YKEEGDEPGSSLLANSPLMEDAPQRLRWQAHLE FTHNHDVGDLTWDKIAVSLPRSEKLRSLVLAGIP HGMRPQLWMRLSGALQKKRNSELSYREIVKNSS NDETIAAKQIEKDLLRTMPSNACFASMGSIGVPR LRRVLRALAWLYPEIGYCQGTGMVAACLLLFLE EEDAFWMMSAIIEDLLPASYFSTTLLGVQTDQRV LRHLIVQYLPRLDKLLQEHDIELSLITLHWFLTAF ASVVDIKLLLRIWDLFFYEGSRVLFQLTLGMLHL KEEELIQSENSASIFNTLSDIPSQMEDAELLLGVA MRLAGSLTDVAVETQRRKHLAYLIADQGQLLGA GTLTNLSQVVRRRTQRRKSTITALLFGEDDLEAL KAKNIKQTELVADLREAILRVARHFQCTDPKNCS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VVSRQLPGLLPNTALTPPTPLVGLCSLWQELTPD YSMESHQRDHENYVACSRSHRRRAKALLDFERH DDDELGFRKNDIITIVSQKDEHCWVGELNGLRG WFPAKFVEVLDERSKEYSIAGDDSVTEGVTDLV RGTLCPALKALFEHGLKKPSLLGGACHPWLFIEE AAGREVERDFASVYSRLVLCKTFRLDEDGKVLT PEELLYRAVQSVNVTHDAVHAQMDVKLRSLICV GLNEQVLHLWLEVLCSSLPTVEKWYQPWSFLRS PGWVQIKCELRVLCCFAFSLSQDWELPAKREAQ QPLKEGVRDMLVKHHLFSWDVDG
3139	A	110	2499	QDRRLLRLELQKTCQPTSTMSGSHTPACGPFSAL TPSIWPQEILAKYTQKEESAEQPEFYYDEFGFRV YKEEGDEPGSSLLANSPLMEDAPQRLRWQAHLE FTHNHDVGDLTWDKIAVSLPRSEKLRSLVLAGIP HGMRPQLWMRLSGALQKKRNSELSYREIVKNSS NDETIAAKQIEKDLLRTMPSNACFASMGSIGVPR LRRVLRALAWLYPEIGYCQGTGMVAACLLLFLE EEDAFWMMSAIIEDLLPASYFSTTLLGVQTDQRV LRHLIVQYLPRLDKLLQEHDIELSLITLHWFLTAF ASVVDIKLLLRIWDLFFYEGSRVLFQLTLGMLHL KEEELIQSENSASIFNTLSDIPSQMEDAELLLGVA MRLAGSLTDVAVETQRRKHLAYLIADQGQLLGA GTLTNLSQVVRRRTQRRKSTITALLFGEDDLEAL KAKNIKQTELVADLREAILRVARHFQCTDPKNCS VVSRQLPGLLPNTALTPPTPLVGLCSLWQELTPD YSMESHQRDHENYVACSRSHRRRAKALLDFERH DDDELGFRKNDIITIVSQKDEHCWVGELNGLRG WFPAKFVEVLDERSKEYSIAGDDSVTEGVTDLV RGTLCPALKALFEHGLKKPSLLGGACHPWLFIEE AAGREVERDFASVYSRLVLCKTFRLDEDGKVLT PEELLYRAVQSVNVTHDAVHAQMDVKLRSLICV GLNEQVLHLWLEVLCSSLPTVEKWYQPWSFLRS PGWVQIKCELRVLCCFAFSLSQDWELPAKREAQ QPLKEGVRDMLVKHHLFSWDVDG
3140	A	1	4939	SAALGASLAIPRPGLPGVHGRGPGTLSGRAMEG AEPRARPERLAEAETRAADGGRLVEVQLSGGAP WGFTLKGGREHGEPLVITKIEEGSKAAAVDKLL AGDEIVGINDIGLSGFRQEAICLVKGSHKTLKLV VKRRSELGWRPHSWHATKFSDSHPELAASPFTST SGCPSWSGRHHASSSSHDLSSSWEQTNLQRTLD HFSSLGSVDSLDHPSSRLSVAKSNSSIDHLGSHSK RDSAYGSFSTSSSTPDHTLSKADTSSAENILYTVG LWEAPRQGGRQAQAAGDPQGSEEKLSCFPPRVP GDSGKGPRPEYNAEPKLAAPGRSNFGPVWYVPD KKKAPSSPPPPPPPLRSDSFAATKSHEKAQGPVFS EAAAAQHFTALAQAQPRGDRRPELTDRPWRSAH PGSLGKGSGGPGCPQEAHADGSWPPSKDGASSR LQASLSSSDVRFPQSPHSGRHPPLYSDHSPLCADS LGQEPGAASFQNDSPPQVRGLSSCDQKLGSGWQ GPRPCVQGDLQAAQLWAGCWPSDTALGALESL PPPTVGQSPRHHLPQPEGPPDARETGRCYPLDKG AEGCSAGAQEPPRASRAEKASQRLAASITWADG ESSRICPQETPLLHSLTQEGKRRPESSPEDSATRPP PFDAHVGKPTRRSDRFATTLRNEIQMHRAKLQK SRSTVALTAAGEAEDGTGRWRAGLGGGTQEGPL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AGTYKDHLKEAQARVLRATSFKRRDLDPNPGDLYPESLEHRMGDPDTVPHFWEAGLAQPPSSTSGGPHPPRIGGRRRFTAEQKLKSYSEPEKMNEVGLTRGYSPHQHPRTSEDTVGTFADRWKFFEETSKPVPQRPAQKQALHGIPRDKPERPRTAGRTCEGTEPWSRTTSLGDSLNAHSAAEKAGTSDLPRRLGTFAEYQASWKEQRKPLEARSSGRCHSADDILDVSLDPQERPQHVHGRSRSSPSTDHYKQEASVELRRQAGDPGEPREELPSAVRAEEGQSTPRQADAQCREGSPGSQQHPPSQKAPNPPTFSELSHCRGAPELPREGRGRAGTLPRDYRYSEESTPADLGPRAQSPGSPLHARGQDSWPVSSALLSKRPAPQRPPPPKREPRRYRATDGAPADAPVGVLGRPFPTPSPASLDVYVARLSLSHSPSVFSSAQPQDTPKATVCERGSQHVSGDASRPLPEALLPPKQQHLRLQTATMETSRSPSPQFAPQKLTDKPPLLIQDEDSTRIERVMDNNTTVKMVPIKIVHSESQPEKESRQSLACPAEPPALPHGLEKDQIKTLSTSEQFYSRFCLYTRQGAEPEAPHRAQPAEPQPLGTQVPPEKDRCTSPPGLSYMKAKEKTVEDLKSEELAREIVGKDKSLADILDPSVKIKTTMDLMEGIFPKDEHLLEEAQQRRKLLPKIPSPRSTEERKEEPSVPAAVSLATNSTYYSTSAPKAELLIKMKDLQEQQEHEEDSGSDLDHDLSVKKQELIESISRKLQVLREARESLLEDVQANTVLGAEVEAIVKGVCKPSEFDKFRMFIGDLDKVVNLLLSLSGRLARVENALNNLDDGASPGDRQSLLEKQRVLIQQHEDAKELKENLDRRERIVFDILANYLSEESLADYEHFVKMKSALIIEQRELEDKIHLGEEQLKCLLDSLQPERGK
3141	A	97	1894	SPRGATMETPPLPPACTKQGHQKPLDSKDDNTEKHCPVTVNPWHMKKAFKVMNELRSQNLLCDVTIVAEDMEISAHRVVLAACSPYFHAMFTGEMSESRAKRVRIKEVDGWTLRMLIDYVYTAEIQVTEENVQVLLPAAGLLQLQDVKKTCCEFLESQLHPVNCLGIRAFADMHACTDLLNKANTYAEQHFADVVLSEEFLNLGIEQVCSLISSDKLTISSEEKVFEAVIAWVNHDKDVRQEFMARLMEHVRLPLLPREYLVQRVEEEALVKNSSACKNYLIEAMKYHLLPTEQRILMKSVRTRLRTPMNLPKLMVVVGGQAPKAIRSAECYDFKEQRWHQVAELPSRRCRAGMVYLAGLVFAVGGFNGSLRVRTVDSYDPVKDQWTSVANMRDRRSTLGAAVLNGLLYAVGGFDGSTGLSSVEAYNIKSNEWFHVAPMNTRRSSVGVGVVGGLLYAVGGYDGASRQYLSTVECYNATTNEWTYIAEMSTRRSGAGVGVLNNLLYAVGGHDGPLVRKSVEVYDPTTNAWRQVADMNMCRRNAGVCAVNGLLYVVGGDDGSCNLASVEYYNPTTDKWTVVSSCMSTGRSYAGVTVIDKPL
3142	A	1211	1311	FSNLTTEKVAHAKEENLSMHQMLDQTLLELNNM
3143	A	1809	1041	SEELDREKKLKEDSPRKTPNKESGVPSLPVSLTSIKEEPKEAKHPDSQSMEESKLKNDDRKTPVNWKDSRGTRVAVSSPMSQHQSYIQYLHAYPYPQMYDPSHPAYRAVSPVLMHSYPGAYLSPGFHYPVYGKMSGREETEKVNTSPSVNTKTTTESKALDLLQQHANQYRSKSPAPVEKATAERERERERERDRHSPFG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QRHLHTHHHTHVGMGYPLIPGQYDPFQGLTSAA LVASQQVAAQASASGMFPGQRR
3144	A	78	604	SVSGIVLDLLPYLHFLSNMNLDGSAQDPEKREYS SVCVGREDDIKKSERMTAVVHDREVVIFYHKGE YHAMDIRCYHSGGPLHLGDIEDFDGRPCIVCPW HKYKITLATGEGLYQSINPKDPSAKPKWCSKGIK QRIHTVTVDNGNIYVTLSNEPFKCDSDFYATGDF KVIKSSS
3145	A	2	333	RNSLLLPPLHLDNSTPAKMSCQQNQQQCQPPPK CPSPKCPPKSPVQCLPPASSGCAPSSGGCGPSSEG GCFLNHHRRHHRCRRQRPNSCDRGSGQQGGGS GCGHGSGGCC
3146	A	3	1151	VCTALQEFGTRSTLLRCLDSGFRPGASRGLVGSW AAMESTLGAGIVIAEALQNQLAWLENVWLWITF LGDPKILFLFYFPAAYYASRRVGIAVLWISLITEW LNLIFKWFLFGDRPFWWVHESGYYSQAPAQVHQ FPSSCETGPGSPSGHCMITGAALWPIMTALSSQV ATRARSRWVRVMPSLAYCTFLLAVGLSRIFILAH FPHQVLAGLITGAVLGWLMTPRVPMERELSFYG LTALALMLGTSLIYWTLFTLGLDLSWSISLAFKW CERPEWIHVDSRPFASLSRDSGAALGLGIALHSPC YAQVRRAQLGNGQKIACLVLAMGLLGPLDWLG HPPQISLFYIFNFLKYTLWPCLVLALVPWAVHMF SAQEAPPIHSS
3147	A	1437	594	RSFSLSFSLLSPSEMMALGAAGATRVFVAMVAA ALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAA GHPGSAVSAAPGILYPGGNKYQTIDNYQPYPCAE DEECGTDEYCASPTRGGDAGVQICLACRKRRKR CMRHAMCCPGNYCKNGICVSSDQNHFRGEIEETI TESFGNDHSTLDGYSRRTTLSSKMYHTKGQEGS VCLRSSDCASGLCCARHFWSKICKPVLKEGQVC TKHRRKGSHGLEIFQRCYCGEGLSCRIQKDHHQ ASNSSRLHTCQRH
3148	A	1	1562	MSTLYDIRAHKAQLLRFFASSDSNKALEQRRTLH TPKLEHLDRVLYEWFLGKRSEGVPVSGPMLIEK AKDFYEQMQLTEPCVFSGGWLWRFKARHGIKK LDASSEKQSADHQAAEQFCAFFRSLAAEHGLSA EQVYNADETGLFWRCLPNPTPEGGAVPGPKQGK DRLTVLMCANATGSHRLKPLAIGKCSGPRAFKGI QHLPVAYKAQGNAWVDKEIFSDWFHHIFVPSVR EHFRTIGLPEDSKAVLLLDSSRAHPQEAELVSSN VFTIFLPASVASLVQPMEQGIRRDFMRNFINPPVP LQGPHARYNMNDAIFSVACAWNAVPSHVFRRA WRKLWPSVAFAEGSSSEEELEAECFPVKPHNKSF AHILELVKEGSSCPGQLRQRQAASWGVAGREAE GGRPPAATSPAEVVWSSEKTPKADQDGRGDPGE GEEVAWEQAAVAFDAVLRFAERQPCFSAQEVG QLRALRAVFRSQQQVRRRRGALGAVVKVEALQ EGPGGCGATAQSPLPCSSTAGDN
3149	A	132	4125	VAVMISTAPLYSGVHNWTSSDRIRMCGINEERRA PLSDEESTTGDCQHFGSQEFCVSSSFSKVELTAV GSGSNARGADPDGSATEKLGHKSEDKPDDPQPK MDYAGNVAEAEGLLVPLSSPGDGLKLPASDSAE ASNSRADCSWTPLNTQMSKQVDCSPAGVKALDS RQGVGEKNTFILATLGTGVPVEGTLPLVTTNFSP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LPAPICPPAPSSASVPHSVPDAFQAPVPPSAPTLVL APVPTPVLAPMPASTPPAAPAPPSVPMPTPTPSSG PPSTPTLIPAFAPTPVPAPTPAPIFTPAPTPMPAATP AAIPTSAPIPASFSLSRVCFPAAQAPAMQKVPLSF QPGTVLTPSQPLVYIPPPSCGQPLSVATLPTTLGV SSTLTLPVLPSYLQDRCLPGVLASPELRSYPYAFS VARPLTSDSKLVSLEVNRLPCTSPSGSTTTQPAPD GVPGPLADTSLVTASAKVLPTPQPLLPAPSGSSAP PHPAKMPSGTEQQTEGTSVTFSPLKSPPQLEREM ASPPECSEMPLDLSSKSNRQKLPLPNQRKTPPMP VLTPVHTSSKALLSTVLSRSQRTTQAAGGNVTSC LGSTSSPFVIFPEIVRNGDPSTWVKNSTALISTIPG TYVGVANPVPASLLLNKDPNLGLNRDPRHLPKQ EPISIIDQGEPKGTGATCGKKGSQAGAEGQPSTV KRYTPARIAPGLPGCQTKELSLWKPTGPANIYPR CSVNGKPTSTQVLPVGWSPYHQASLLSIGISSAG QLTPSQGAPIRPTSVVSEFSGVPSLSSSEAVHGLP EGQPRPGGSFVPEQDPVTKNKTCRIAAKPYEEQV NPVLLTLSPQTGTLALSVQPSGGDIRMNQGPEES ESHLCSDSTPKMEGPQGACGLKLAGDTKPKNQV LATYMSHELVLATPQNLPKMPELPLLPHDSHPKE LILDVVPSSRRGSSTERPQLGSQVDLGRVKMEKV DGDVVFNLATCFRADGLPVAPQRGQAEVRAKA GQARVKQESVGVFACKNKWQPDDVTESLPPKK MKCGKEKDSEEQQLQPQAKAVVRSSHRPKCRK LPSDPQESTKKSPRGASDSGKEHNGVRGKHKHR KPTKPESQSPGKRADSHEEGSLEKKAKSSFRDFIP VVLSTRTRSQSDLKARKQKTSSSQSLEHRLRNRN LLLPNKVQGISDSPNGFLPNNLEEPACLENSEKPS GKRKCKTKHMATVSEEAKGKGRWSQQKTRSPK SPTPVKPTEPCTPSKSRSASSEEASESPTARQIPPE ARRLIVNKNAGETLLQRAARLGYKDVVLYCLQK DSEDVNHRDNAGYTALHEACSRGWTDILNILLE HGA
3150	A	3	2795	SLRMHNLSILVRQIKFYYQETLQQLIMMSLPNVLI IGKNPFSEQGTEEVKKLLLLLLGCAVQCQKKEEF IERIQGLDFDTKAAVAAHIQEVTHNQENVFDLQ WMEVTDMSQEDIEPLLKNMALHLKRLIDERDEH SETIIELSEERDGLHFLPHASSSAQSPCGSPGMKR TESRQHLSVELADAKAKIRRLRQELEEKTEQLLD CKQELEQMEIELKRLQQENMNLLSDARSARMYR DELDALREKAVRVDKLESEVSRYKERLHDIEFY KARVEELKEDNQVLLETKTMLEDQLEGTRARSD KLHELEKENLQLKAKLHDMEMERDMDRKKIEE LMEENMTLEMAQKQSMDESLHLGWELEQISRTS ELSEAPQKSLGHEVNELTSSRLLKLEMENQSLTK TVEELRTTVDSVEGNASKILKMEKENQRLSKKV EILENEIVQEKQSLQNCQNLSKDLMKEKAQLEKT IETLRENSERQIKILEQENEHLNQTVSSLRQRSQIS AEARVKDIEKENKILHESIKETSSKLSKIEFEKRQI KKELEHYKEKGERAEELENELHHLEKENELLQK KITNLKITCEKIEALEQENSELERENRKLKKTLDS FKNLTFQLESLEKENSQLDEENLELRRNVESLKC ASMKMAQLQLENKELESEKEQLKKGLELLKASF KKTERLEVSYQGLDIENQRLQKTLENSNKKIQQL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ESELQDLEMENQTLQKNLEELKISSKRLEQLEKE NKSLEQETSQLEKDKKQLEKENKRLRQQAEIKD TTLEENNVKIGNLEKENKTLSKEIGIYKESCVRLE ELEKENKELVKRATIDIKTLVTLREDLVSEKLKT QQMNNDLEKLTHELEKIGLNKERLLHDEQSTDD SRYKLLESKLESTLKKSLEIKEEKIAALEARLEES TNYNQQLRQELKTVKKK
3151	A	2	2515	GFWLHLTLLGASLPAALGWMDPGTSRGPDVGV GESQAEEPRSFEVTRREGLSSHNELLASCGKKFC SRGSRCVLSRKTGEPECQCLEACRPSYVPVCGSD GRFYENHCKLHRAACLLGKRITVIHSKDCFLKGD TCTMAGYARLKNVLLALQTRLQPLQEGDSRQDP ASQKRLLVESLFRDLDADGNGHLSSSELAQHVL KKQDLDEDLLGCSPGDLLRFDDYNSDSSLTLREF YMAFQVVQLSLAPEDRVSVTTVTVGLSTVLTCA VHGDLRPPIIWKRNGLTLNFLDLEDINDFGEDDS LYITKVTTIHMGNYTCHASGHEQLFQTHVLQVN VPPVIRVYPESQAQEPGVAASLRCHAEGIPMPRIT WLKNGVDVSTQMSKQLSLLANGSELHISSVRYE DTGAYTCIAKNEVGVDEDISSLFIEDSARKTLANI LWREEGLSVGNMFYVFSDDGIIVIHPVDCEIQRH LKPTEKIFMSYEEICPQREKNATQPCQWVSAVNV RNRYIYVAQPALSRVLVVDIQAHKVLQSIGVDPL PAKLSYDKSHDQVWVLSWGDVHKSRPSLQVITE ASTGQSQHLIRTPFAGVDDFFIPPTNLIINHIRFGFI FNKSDPAVHKVDLETMMPLKTIGLHHHGCVPQA MAHTHLGGYFFIQCRQDSPASAARQLLVDSVTD SVLGPNGDVTGTPHTSPDGRFIVSAAADSPWLHV QEITVRGEIQTLYDLQINSGISDLAFQRSFTESNQ YNIYAALHTEPDLLFLELSTGKVGMLKNLKEPPA GPAQPWGGTHRIMRDSGLFGQYLLTPARESLFLI NGRQNTLRCEVSGIKGGTTVVWVGEV
3152	A	1	2645	GAGWQVSLTGRWSPGREAGAGEVRQDPGSTAA SPSSCDADLSARMARGERRRRAVPAEGVRTAER AARGGPGRRDGRGGGPRSTAGGVALAVVVLSL ALGMSGRWVLAWYRARRAVTLHSAPAVLPADS SSPAVAPDLFWGTYRPHVYFGMKTRSPKPLLTG LMWAQQGTTPGTPKLRHTCEQGDGVGPYGWEF HDGLSFGRQHIQDGALRLTTEFVKRPGGQHGGD WSWRVTVEPQDSGTSALPLVSLFFYVVTDGKEV LLPEVGAKGQLKFISGHTSELGDFRFTLLPPTSPG DTAPKYGSYNVFWTSNPGLPLLTEMVKSRLNSW FQHRPPGASPERYLGLPGSLKWEDRGPSGQGQG QFLIQQVTLKIPISIEFVFESGSAQAGGNQALPRLA GSLLTQALESHAEGFRERFEKTFQLKEKGLSSGE QVLGQAALSGLLGGIGYFYGQGLVLPDIGVEGSE QKVDPALFPPVPLFTAVPSRSFFPRGFLWDEGFH QLVVQRWDPSLTREALGHWLGLLNADGWIGRE QILGDEARARVPPEFLVQRAVHANPPTLLLPVAH MLEVGDPDDLAFLRKALPRLHAWFSWLHQSQA GPLPLSYRWRGRDPALPTLLNPKTLPSGLDDYPR ASHPSVTERHLDLRCWVALGARVLTRLAEHLGE AEVAAELGPLAASLEAAESLDELHWAPELGVFA DFGNHTKAVQLKPRPPQGLVRVVGRPQPLQYV DALGYVSLFPLLLRLLDPTSSRLGPLLDILADSRH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LWSPFGLRSLAASSSFYGQRNSEHDPPYWRGAVWLNVNYLALGALHHYGHLEGPHQARAAKLHGELRANVVGNVWRQYQATGFLWEQYSDRDGRGMGCRPFHGWTSLVLLAMAEDY
3153	A	1	4312	MVIKTDELPAAAPADSAREHGSQAGGKGRPGAAAVLLADLERDARQGECALPGAAMAGLAPLKPEASRSSSPGPTGCIRARVAAEAGTRNPGNAGAELESWLPCCHGHPETPEPRGGQLPTAPELPSVMLLNGDCPESLKKEAAAAEPPRENGLDEAGPGDETTGQEVIVIQDTGFSVKILAPGIEPFSLQVSPQEMVQEIHQVLMDREDTCHRTCFSLHLDGNVLDHFSELRSVEGLQEGSVLRVVEEPYTVREARIHVRHVRDLLKSLDPSDAFNGVDCNSLSFLSVFTDGDLGDSGKRKKGLEMDPIDCTPPEYILPGSRERPLCPLQPQNRDWKPLQCLKVLTMSGWNPPPGNRKMHGDLMYLFVITAEDRQVSITASTRGFYLNQSTAYHFNPKPASPRFLSHSLVELLNQISPTFKKNFAVLQKKRVQRHPFERIATPFQVYSWTAPQAEHAMDCVRAEDAYTSRLGYEEHIPGQTRDWNEELQTTRELPRKNLPERLLRERAIFKVHSDFTAAATRGAMAVIDGNVMAINPSEETKMQMFIWNNIFFSLGFDVRDHYKDFGGDVAAYVAPTNDLNGVRTYNAVDVEGLYTLGTVVVDYRGYRVTAQSIIPGILERDQEQSVIYGSIDFGKTVVSHPRYLELLERTSRPLKILRHQVLNDRDEEVELCSSVECKGIIGNDGRHYILDLLRTFPPDLNFLPVPGEELPEECARAGFPRAHRHKLCCLRQELVDAFVEHRYLLFMKLAALQLMQQNASQLETPSSLENGGPSSLESKSEDPPGQEAGSEEEGSSASGLAKVKELAETIAADDGTDPRSREVIRNACKAVGSISSTAFDIRFNPDIFSPGVRFPESCQDEVRDQKQLLKDAAAFLLSCQIPGLVKDCMEHAVLPVDGATLAEVMRQRGINMRYLGKVLELVLRSPARHQLDHVFKIGIGELITRSAKHIFKTYLQGVELSGLSAAISHFLNCFLSSYPNPVAHLPADELVSKKRNKRRKNRPPGAADNTAWAVMTPQELWKNICQEAKNYFDFDLECETVDQAVETYGLQKITLLREISLKTGIQVLLKEYSFDSRHKPAFTEEDVLNIFPVVKHVNPKASDAFHFFQSGQAKVQQGFLKEGCELINEALNLFNNVYGAMHVETCACLRLLARLHYIMGDYAEALSNQQKAVLMSERVMGTEHPNTIQEYMHLALYCFASSQLSTALSLLYRARYLMLLVFGEDHPEMALLDNNIGLVLHGVMEYDLSLRFLENALAVSTKYHGPKALKVALSHHLVARVYESKAEFRSALQHEKEGYTIYKTQLGEDHEKTKESSEYLKCLTQQAVALQRTMNEIYRNGSSANIPPLKFTAPSMASVLEQLNVINGILFIPLSQKDLENLKAEVARRHQLQEASRNRDRAEEPMATEPAPAGAPGDLGSQPPAAKDPSPSVQG
3154	A	416	4082	KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHWSCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRIDTEGTNYEQLVVDAGVSVIMDFHYNEKRIYWVDLERQLLQRVFLNGSRQERVCNIEKNVSGMAINWINEEVIWSNQQEGIITVTDMKGNNSHILLSALKYPANVAVDPVERFIFWSSEVAGSLYRADLDGVGVKALLETSEKITAVSLDVLDKRLFWIQYNREGSNSLICSCDYDGGSVHISKHPTQHNLFAMSLFGDRIFYSTWK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MKTIWIANKHTGKDMVRINLHSSFVPLGELKVV HPLAQPKAEDDTWEPEQKLCKLRKGNCSSTVCG QDLQSHLCMCAEGYALSRDRKYCEGNDWKYCE DVNECAFWNHGCTLGCKNTPGSYYCTCPVGFVL LPDGKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPLSP VSWECDCFPGYDLQLDEKSCAASGPQPFLLFANS QDIRHMHFDGTDYGTLLSQQMGMVYALDHDPV ENKIYFAHTALKWIERANMDGSQRERLIEEGVD VPEGLAVDWIGRRFYWTDRGKSLIGRSDLNGKR SKIITIENISQPRGIAVHPMAKRLFWTDTGINPRIE SSSLQGLGRLVIASSDLIWPSGITIDFLTDKLYWC DAKQSVIEMANLDGSKRRRLTQNDVGHPFAVA VFEDYVWFSDWAMPSVIRVNKRTGKDRVRLQG SMLKPSSLVVVHPLAKPGADPCLYQNGGCEHIC KKRLGTAWCSCREGFMKASDGKTCLALDGHQL LAGGEVDLKNQVTPLDILSKTRVSEDNITESQHM LVAEIMVSDQDDCAPVGCSMYARCISEGEDATC QCLKGFAGDGKLCSDIDECEMGVPVCPPASSKCI NTEGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPDSTP PPHLREDDHHYSVRNSDSECPLSHDGYCLHDGV CMYIEALDKYACNCVVGYIGERCQYRDLKWWE LRHAGHGQQQKVIVVAVCVVVLVMLLLLSLWG AHYYRTQKLLSKNPKNPYEESSRDVRSRRPADT EDGMSSCPQPWFVVIKEHQDLKNGGQPVAGED GQAADGSMQPTSWRQEPQLCGMGTEQGCWIPV SSDKGSCPQVMERSFHMPSYGTQTLEGGVEKPH SLLSANPLWQQRALDPPHQMELTQ
3155	A	533	212	GTSGWYWERLAERRGRLWSREEAMATMENKVI CALVLVSMLALGTLAEAQTETCTVAPRERQNCG FPGVTPSQCANKGCCFDDTVRGVPWCFYPNTID VPPEEECEF
3156	A	2	1585	PRVRAADVAAGAQAVVSAGMAKSNGENGPRAP AAGESLSGTRESLAQGPDAATTDELSSLGSDSEA NGFAERRIDKFGFIVGSQGAEGALEEVPLEVLRQ RESKWLDMLNNWDKWMAKKHKKIRLRCQKGI PPSLRGRAWQYLSGGKVKLQQNPGKFDELDMSP GDPKWLDVIERDLHRQFPFHEMFVSRGGHGQQD LFRVLKAYTLYRPEEGYCQAQAPIAAVLLMHMP AEQAFWCLVQICEKYLPGYYSEKLEAIQLDGEIL FSLLQKVSPVAHKHLSRQKIDPLLYMTEWFMCA FSRTLPWSSVLRVWDMFFCEGVKIIFRVGLVLLK HALGSPEKVKACQGQYETIERLRSLSPKIMQEAF LVQEVVELPVTERQIEREHLLQLRRWQETRGELQ CRSPPRLHGAKAILDAEPGPRPALQPSPSIRLPLD APLPGSKAKPKPPKQAQKEQRKQMKGRGQLEKP PAPNQAMVVAAAGDACPPQHVPPKDSAPKDSAP QDLAPQVSAHHRSQESLTSQESEDTYL
3157	A	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLLRLH HRFRALDRNKKGYLSRMDLQQIGALAVNPLGDR IIESFFPDGSQRVDFPGFVRVLAHFRPVEDEDTET QDPKKPEPLNSRRNKLHYAFQLYDLDRDGKISR HEMLQVLRLMVGVQVTEEQLENIADRTVQEAD EDGDGAVSFVEFTKSLEKMDVEHKMSIRILK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3158	A	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFITYMDNWRQNTTAEQEALQAKVDAENFYYVILYLMVMIGMFSFIIVAILVSTVKSKRREHSNDPYHQYIVEDWQEKYKSQILNLEESKATIHENIGAAGFKMSP
3159	A	3	416	PWGAAELDMGRRDAQLLAALLVLGLCALAGSEKPSPCQCSRLSPHNRTNCGFPGITSDQCFDNGCCFDSSVTGVPWCFHPLPKQESDQCVMEVSDRRNCGYPGISPEECASRKCCFSNFIFEVPWCFFPKSVEDCHY
3160	A	179	409	KPKTKILKMVYYPELFVWVSQEPFPNKDMEGRLPKGRLPVPKEVNRKKNDETNAASLTPLGSSELRSPRISYLHFF
3161	A	683	1186	LSSTGGLHAAACAAAMSLVIPEKFQHILRVLNTNIDGRRKIAFAITAIKGVGRRYAHVVLRKADIDLTKRAGELTEDEVERVITIMQNPRQYKIPDWFLNRQKDVKDGKYSQVLANGLDNKLREDLERLKKIRAHRGLRHFWGLRVRGQHTKTTGRRGRTVGVSKKK
3162	A	1	1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWRVPGRLLLLLLLPALCCLPGAARAAAAAAGAGNRAAVAVAVARADEAEAPFAGQNWLKSYGYLLPYDSRASALHSAKALQSAVSTMQQFYGIPVTGVLDQTTIEWMKKPRCGVPDHPHLSRRRRNKRYALTGQKWRQKHITYSIHNYTPKVGELDTRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDRKEADIMIFFASGFHGDSSPFDGEGGFLAHAYFPGPGIGGDTHFDSDEPWTLGNANHDGNDLFLVAVHELGHALGLEHSSDPSAIMAPFYQYMETHNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTLPVRRIHSPSERKHERQPRPPRPPLGDRPSTPGTKPNICDGNFNTVALFRGEMFVFKDRWFWRLRNNRVQEGYPMQIEQFWKGLPARIDAAYERADGRFVFFKGDKYWVFKEVTVEPGYPHSLGELGSCLPREGIDTALRWEPVGKTYFFKGERYWRYSEERRATDPGYPKPITVWKGIPQAPQGAFISKEGYYTYFYKGRDYWKFDNQKLSVEPGYPRNILRDWMGCNQKEVERRKERRLPQDDVDIMVTINDVPGSVNAVAVVIPCILSLCILVLVYTIFQFKNKTGPQPVTYYKRPVQEWV
3163	A	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLYTRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEKSPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYNNRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQSKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSGHGRHRR
3164	A	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAERTEAPGTPEGPEPERPSPGDGNPRENSPFLNNVEVEQESFFEGKNMALFEEEMDSNPMVSSLLNKLANYTNLSQGVVEHEEDEESRRREAKAPRMGTFIGVYLPCLQNILGVILFLRLTWIVGVAGVLESFLIVAMCCTCTMLTAISMSAIATNGVVPAGGSYYMISRSLGPEFGGAVGLCFYLGTTFAGAMYILGTIEIFLTYISP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GAAIFQAEAAGGEAAAMLHNMRVYGTCTLVLM ALVVFVGVKYVNKLALVFLACVVLSILAIYAGVI KSAFDPPDIPVCLLGNRTLSRRSFDACVKAYGIH NNSATSALWGLFCNGSQPSAACDEYFIQNNVTEI QGIPGAASGVFLENLWSTYAHAGAFVEKKGVPS VPVAEESRASTLPYVLTDIAASFTLLVGIYFPSVT GIMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIYLS CIVLFGACIEGVVLRDKFGEALQGNLVIGMLAW PSPWVIVIGSFFSTCGAGLQTLTGAPRLLQAIARD GIVPFLQVFGHGKANGEPTWALLLTVLICETGILI ASLDSVAPILSMFFLMCYLFVNLACAVQTLLRTP NWRPRFKFYHWTLSFLGMSLCLALMFICSWYYA LSAMLIAGCIYKYIEYRGAEKEWGDGIRGLSLNA ARYALLRVEHGPPHTKNWRPQVLVMLNLDAEQ AMKHPRLLSFTSQLKAGKGLTIVGSVLEGTYLD KHMEAQRAEENIRSLMSTEKTKGFCQLVVSSSLR DGMSHLIQSAGLGGLKHNTVLMAWPASWKQED NPFSWKNFVDTVRDTTAAHQALLVAKNVDSFPQ NQERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMFLY HLRISAEVEVVEMVENDISAFTYERTLMMEQRS QMLKQMQLSKNEQEREAQLIHDRNTASHTAAA ARTQAPPTPDKVQMTWTREKLIAEKYRSRDTSL SGFKDLFSMKPDQSNVRRMHTAVKLNGVVLNK SQDAQLVLLNMPGPPKNRQGDENYMEFLEVLTE GLNRVLLVRGGGREVITIYS
3165	A	3	2681	GRGARGGSGAGALRGCRGYLQKLSGKGPSRGY RSRWFVFDARRCYLYYFKSPQDALPLGHLDIAD ACFSYQGPDEAAEPGTEPPAHFQVHSAGAVTVL KAPNRQLMTYWLQELQQKRWEYCNSLDMVKW DSRTSPTPGDFPKGLVARDNTDLIYPHPNASAEK ARNVLAVETVPGELVGEQAANQPAPGHPNSINF YSLKQWGNELKNSMSSFRPGRGHNDSRRTVFYT NEEWELLDPTPKDLEESIVQEEKKKLTPEGNKGV TGSGFPFDFGRNPYKGKRPLKDIIGSYKNRHSSG DPSSEGTSGSGSVSIRKPASEMQLQVQSQQEELE QLKKDLSSQKELVRLLQQTVRSSQYDKYFTSSRL CEGVPKDTLELLHQKDDQILGLTSQLERFSLEKE SLQQEVRTLKSKVGELNEQLGMLMETIQAKDEV IIKLSEGEGNGPPPTVAPSSPSVVPVARDQLELDR LKDNLQGYKTQNKFLNKEILELSALRRNPERRER DLMARNSSLEAKLCQIESKYLILLQEMKTPVCSE DQGPTREVIAQLLEDALQVESQEQPEQAFVKPHL VSEYDIYGFRTVPEDDEEEKLVAKVRALDLKTL YLTENQEVSTGVKWENYFASTVNREMMCSPEL KNLIRAGIPHEHRSKVWKWCVDRHTRKFKDNTE PGHFQTLLQKALEKQNPASKQIELDLLRTLPNNK HYSCPTSEGIQKLRNVLLAFSWRNPDIGYCQGLN RLVAVALLYLEQEDAFWCLVTIVEVFMPRDYYT KTLLGSQVDQRVFRDLMSEKLPRLHGHFEQYKV DYTLITFNWFLVVFVDSVVSDILFKIWDSFLYEGP KVIFRFALALFKYKEEEILKLQDSMSIFKYLRYFT RTILDARSGTDAPTTWRKSGWS
3166	A	10	4070	FPGPTISSNSQLYRASALFETIRHEAQLSTDYKLS LFDLQTSSYQALQRVLVSLGHHDEALAVAERGR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TRAFADLLVERQTGQQDSDPYSPVTIDQILEMVN GQRGLVLYYSLAAGYLYSWLLAPGAGIVKFHEH YLGENTVENSSDFQASSSVTLPTATGSALEQHIAS VREALGVESHYSRACASSETESEAGDIMDQQFEE MNNKLNSVTDPTGFLRMVRRNNLFNRSCQSMTS LFSNTVSPTQDGTSSLPRRQSSFAKPPLRALYDLL IAPMEGGLMHSSGPVGRHRQLILVLEGELYLIPF ALLKGSSSNEYLYERFGLLAVPSIRSLSVQSKSHL RKNPPTYSSSTSMAAVIGNPKLPSAVMDRWLWG PMPSAEEEAYMVSELLGCQPLVGSVATKERVMS ALTQAECVHFATHISWKLSALVLTPSMDGNPASS KSSFGHPYTIPESLRVQDDASDGESISDCPPLQEL LLTAADVLDLQLPVKLVVLGSSQESNSKVAADG VIALTRAFLAAGAQCVLVSLWPVPVAAFKMFIH AFYSSLLNGLKASAALGEAMKVVQSSKAFSHPS NWAGFMLIGSDVKLNSPSSLIGQALTEILQHPER ARDALRVLLHLVEKSLQRIQNGQRNAMYTSQQS VENKVGGIPGWQALLTAVGFRLDPPTSGLPAAV FFPTSDPGDRLQQCSSTLQSLLGLPNPALQALCK LITASETGEQLISRAVKNMVGMLHQVLVQLQAG EKEQDLASAPIQVSISVQLWRLPGCHEFLAALGF VLCEVGQEEVILKTGKQANRRTVHFALQSLLSLF DSTELPKRLSLDSSSSLESLASAQSVSNALPLGYQ QPPFSPTGADSIASDAISVYSLSSIASSMSFVSKPE GGSEGGGPGGRQDHDRSKNAYLQRSTLPRSQLP PQTRPAGNKDEEEYEGFSIISNEPLATYQENRNTC FSPDHKQPQPGTAGGMRVSVSSKGSISTPNSPVK MTLIPSPNSPFQKVGKLASSDTGESDQSSTETDST VKSQEESNPKLDPQELAQKILEETQSHLIAVERLQ RSGGQVSKSNNPEDGVQAPSSTAVFRASETSAFS RPVLSHQKSQPSPVTVKPKPPARSSSLPKVSSGYS SPTTSEMSIKDSPSQHSGRPSPGCDSQTSQLDQPL FKLKYPSSPYSAHISKSPRNMSPSSGHQSPAGSAP SPALSYSSAGSARSSPADAPDIDKLKMAAIDEKV QAVHNLKMFWQSTPQHSTGPMKIFRGAPGTMTS KRDVLSLLNLSPRPNKKEEGVDKLELKELSLQQH DGAPPKAPPNGHWRTETTSLGSLPLPAGPPATAP ARPLRLPSGNGYKFLSPGRFFPSSKC
3167	A	1	762	AARRRQKGKEENMMMDLFETGSYFFYLDGENV TLQPLEVAEGSPLYPGSDGTLSPCQDQMPPEAGS DSSGEEHVLAPPGLQPPHCPGQCLIWACKTCKRK SAPTDRRKAATLRERRRLKKINEAFEALKRRTVA NPNQRLPKVEILRSAISYIERLQDLLHRLDQQEK MQELGVDPFSYRPKQENLEGADFLRTCSSQWPS VSDHSRGLVITAKEGGASIDSSASSSLRCLSSIVDS ISSEERKLPCVEEVVEK
3168	A	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHVRR KIMSSPLSKELRQKYNVRSMPIRKDDEVQVVRG HYKGQQIGKVVQVYRKKYVIYIERVQREKANGT TVHVGIHPSKVVITRLKLDKDRKKILERKAKSRQ VGKEKGKYKEELIEKMQE
3169	A	156	3168	GPGGAISLSVEAKAGADLLVKGKQARMDIYDTQ TLGVVVFGGFMVVSAIGIFLVSTFSMKETSYEEA LANQRKEMAKTHHQKVEKKKKEKTVEKKGKT KKKEEKPNGKIPDHDPAPNVTVLLREPVRAPAV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AVAPTPVQPPIIVAPVATVPAMPQEKLASSPKDK KKKEKKVAKVEPAVSSVVNSIQVLTSKAAILETA PKEGRNTDVAQSPEAPKQEAPAKKKSGSKKKGP PDADGPLYLPYKTLVSTVGSMVFNEGEAQRLIEI LSEKAGIIQDTWHKATQKGDPVAILKRQLEEKEK LLATEQEDAAVAKSKLRELNKEMAAEKAKAAA GEAKVKKQLVAREQEITAVQARMQASYREHVK EVQQLQGKIRTLQEQLENGPNTQLARLQQENSIL RDALNQATSQVESKQNAELAKLRQELSKVSKEL VEKSEAVRQDEQQRKALEAKAAAFEKQVLQLQ ASHRESEEALQKRLDEVSRELCHTQSSHASLRAD AEKAQEQQQMAELHSKLQSSEAEVRSKCEELS GLHGQLQEARAENSQLTERIRSIEALLEAGQARD AQDVQASQAEADQQQTRLKELESQVSGLEKEAI ELREAVEQQKVKNNDLREKNWKAMEALATAEQ ACKEKLHSLTQAKEESEKQLCLIEAQTMEALLAL LPELSVLAQQNYTEWLQDLKEKGPTLLKHPPAP AEPSSDLASKLREAEETQSTLQAECDQYRSILAET EGMLRDLQKSVEEEEQVWRAKVGAAEEELQKS RVTVKHLEEIVEKLKGELESSDQVREHTSHLEAE LEKHMAAASAECQNYAKEVAGLRQLLLESQSQL DAAKSEAQKQSDELALVRQQLSEMKSHVEDGDI AGAPASSPEAPPAEQDPVQLKTQLEWTEAILEDE QTQRQKLTAEFEEAQTSACRLQEELEKLRTAGPL ESSETEEASQLKERLEKEKKLTSDLGRAATRLQE LLKTTQEQLAREKDTVKKLQEQLEKAEDGSSSK EGTSV
3170	A	6730	4027	THASEKYSYGHLPTHSITAHPMVTIRISDRQRLIQ PYIHNYSWLLFAALALYSAHLASAEDVDGEKLD PQTRSSATTLRSQCMQLVGDCLMKAHQGKGLK ALALLGVLPDGDSSLEDHALPVTVPTGASEEQLE KKAVQGAELSEAGNGKRAVHEEIRPVDFKQRNK ADKGVSLSKDPSCQTQISDSPADASPPTGLPDAE DSEVSSQKPIEEKAVTPSPEQVFAECSQKRILGLL AAMLPPLKSGPTVPLIDLEHVLPLMFQVVISNAG HLNETYHLTLGLLGQLIIRLLPAEVDAAVIKVLSA KHNLFAAGDSSIVPDGWKTTHLLFSLGAVCLDS RVGLDWACSMAEILRSLNSAPLWRDVIATFTDH CIKQLPFQLKHTNIFTLLVLVGFPQVLCVGTRCV YMDNANEPHNVIILKHFTEKNRAVIVDVKTRKR KTVKDYQLVQKGGGQECGDSRAQLSQYSQHFA FIASHLLQSSMDSHCPEAVEATWVLSLALKGLY KTLKAHGFEEIRATFLQTDLLKLLVKKCSKGTGF SKTWLLRDLEILSIMLYSSKKEINALAEHGDLEL DERGDREEEVERPVSSPGDPEQKKLDPLEGLDEP TRICFLMAHDALNAPLHILRAIYELQMKKTDYFF LEVQKRFDGDELTTDERIRSLAQRWQPSKSLRLE EQSAKAVDTDMIILPCLSRPARCDQATAESNPVT QKLISSTESELQQSYAKQRRSKSAALLHKELNCK SKRAVRDYLFRVNEATAVLYARHVLASLLAEWP SHVPVSEDILELSGPAHMTYILDMFMQLEEKHE WEKVVMQTELVLTHQVLPLPHRLPPVSASWSEA TCVAVQLPDRCECSKGRVTVSSPKDWASEELRG PERDFQLNQKALSPSSQFPSAEILRHIR
3171	A	557	89	GTRAGPVKDREAFQRLNFLYQAAHCVLAQDPEN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QALARFYCYTERTIAKRLVLRRDPSVKRTLCRGC SSLLVPGLTCTQRQRRCRGQRWTVQTCLTCQRS QRFLNDPGHLLWGDRPEAQLGSQADSKPLQPLP NTAHSISDRLPEEKMQTQGSSNQ
3172	A	2	496	FRRAGAGRGRRRGEVTSPLSPEPLAFQSLATSRR PEPQTTQTVRSSALPAPPASPMSQYAPSPDFKRA LDSSPEANTEDDKTEEDVPMPKNYLWLTIVSCFC PAYPINIVALVFSIMSLNSYNDGDYEGARRLGRN AKWVAIASIIIGLLIIGISCAVHFTRNA
3173	A	2	4048	FRSGGCRRRAWTSRWPQRRRSPESCEAPLSAPL WGPQRGLPGREPLRSRSASAIALRTIGHILALLLR LLHLGLGSGGCREDVPPSGRGKKEEKMKKHRRA LALVSCLFLCSLVWLPSWRVCCKESSSASASSYY SQDDNCALENEDVQFQKKDEREGPINAESLGKS GSNLPISPKEHKLKDDSIVDVQNTESKKLSPPVVE TLPTVDLHEESSNAVVDSETVENISSSSTSEITPIS KLDEIEKSGTIPIAKPSETEQSETDCDVGEALDAS APIEQPSFVSPPDSLVGQHIENVSSSHGKGKITKSE FESKVSASEQGGGDPKSALNASDNLKNESSDYT KPGDIDPTSVASPKDPEDIPTFDEWKKKVMEVEK EKSQSMHASSNGGSHATKKVQKNRNNYASVEC GAKILAANPEAKSTSAILIENMDLYMLNPCSTKI WFVIELCEPIQVKQLDIANYELFSSTPKDFLVSISD RYPTNKWIKLGTFHGRDERNVQSFPLDEQMYAK YVKMFIKYIKVELLSHFGSEHFCPLSLIRVFGTSM VEEYEEIADSQYHSERQELFDEDYDYPLDYNTGE DKSSKNLLGSATNAILNMVNIAANILGAKTEDLT EGNKSISENATATAAPKMPESTPVSTPVPSPEYVT TEVHTHDMEPSTPDTPKESPIVQLVQEEEEEASPS TVTLLGSGEQEDESSPWFESETQIFCSELTTICCIS SFSEYIYKWCSVRVALYRQRSRTALSKGKDYLV LAQPPLLLPAESVDVSVLQPLSGELENTNIEREAE TVVLGDLSSSMHQDDLVNHTVDAVELEPSHSQT LSQSLLLDITPEINPLPKIEVSESVEYEAGHIPSPVI PQESSVEIDNETEQKSESFSSIEKPSITYETNKVNE LMDNIIKEDVNSMQIFTKLSETIVPPINTATVPDN EDGEAKMNIADTAKQTLISVVDSSSLPEVKEEEQ SPEDALLRGLQRTATDFYAELQNSTDLGYANGN LVHGSNQKESVFMRLNNRIKALEVNMSLSGRYL EELSQRYRKQMEEMQKAFNKTIVKLQNTSRIAE EQDQRQTEAIQLLQAQLTNMTQLVSNLSATVAE LKREVSDRQSYLVISLVLCVVLGLMLCMQRCRN TSQFDGDYISKLPKSNQYPSPKRCFSSYDDMNLK RRTSFPLMRSKSLQLTGKEVDPNDLYIVEPLKFSP EKKKKRCKYKIEKIETIKPEEPLHPIANGDIKGRK PFTNQRDFSNMGEVYHSSYKGPPSEGSSETSSQS EESYFCGISACTSLCNGQSQKTKTEKRALKRRRS KVQDQGKLIKTLIQTKSGSLPSLHDIIKGNKEITV GTFGVTAVSGHI
3174	A	485	4668	RKCSKEKASKTPSQKIPTTPCCVLQAGPEPRSLAE RMGADGETVVLKNMLIGVNLILLGSMIKPSECQL EVTTERVQRQSVEEEGGIANYNTSSKEQPVVFNH VYNINVPLDNLCSSGLEASAEQEVSAEDETLAEY MGQTSDHESQVTFTHRINFPKKACPCASSAQVLQ ELLSRIEMLEREVSVLRDQCNANCCQESAATGQL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DYIPHCSGHGNFSFESCGCICNEGWFGKNCSEPY CPLGCSSRGVCVDGQCICDSEYSGDDCSELRCPT DCSSRGLCVDGECVCEEPYTGEDCRELRCPGDCS GKGRCANGTCLCEEGYVGEDCGQRQCLNACSG RGQCEEGLCVCEEGYQGPDCSAVAPPEDLRVAG ISDRSIELEWDGPMAVTEYVISYQPTALGGLQLQ QRVPGDWSGVTITELEPGLTYNISVYAVISNILSL PITAKVATHLSTPQGLQFKTITETTVEVQWEPFSF SFDGWEISFIPKNNEGGVIAQVPSDVTSFNQTGLK PGEEYIVNVVALKEQARSPPTSASVSTVIDGPTQI LVRDVSDTVAFVEWIPPRAKVDFILLKYGLVGGE GGRTTFRLQPPLSQYSVQALRPGSRYEVSVSAVR GTNESDSATTQFTTEIDAPKNLRVGSRTATSLDL EWDNSEAEVQEYKVVYITLAGEQYHEVLVPRGI GPTTRATLTDLVPGTEYGVGISAVMNSQQSVPAT MNARTELDSPRDLMVTASSETSISLIWTKASGPID HYRITFTPSSGIASEVTVPKDRTSYTLTDLEPGAE YIISVTAERGRQQSLESTVDAFTGFRPISHLHFSH VTSSSVNITWSDPSPPADRLILNYSPRDEEEEMME VSLDATKRHAVLMGLQPATEYIVNLVAVHGTVT SEPIVGSITTGIDPPKDITISNVTKDSVMVSWSPPV ASFDYYRVSYRPTQVGRLDSSVVPNTVTEFTITR LNPATEYEISLNSVRGREESERICTLVHTAMDNP VDLIATNITPTEALLQWKAPVGEVENYVIVLTHF AVAGETILVDGVSEEFRLVDLLPSTHYTATMYAT NGPLTSGTISTNFSTLLDPPANLTASEVTRQSALIS WQPPRAEIENYVLTYKSTDGSRKELIVDAEDTWI RLEGLLENTDYTVLLQAAQDTTWSSITSTAFTTG GRVFPHPQDCAQHLMNGDTLSGVYPIFLNGELS QKLQVYCDMTTDGGGWIVFQRRQNGQTDFFRK WADYRVGFGNVEDEFWLGLDNIHRITSQGRYEL RVDMRDGQEAAFASYDRFSVEDSRNLYKLRIGS YNGTAGDSLSYHQGRPFSTEDRDNDVAVTNCA MSYKGAWWYKNCHRTNLNGKYGESRHSQGIN WYHWKGHEFSIPFVEMKMRPYNHRLMAGRKRQ SLQF
3175	A	2	623	RLQLPACPALSAAHPLALPSFSSQCHRAEARAAA AATAEGTMASGVTVNDEVIKVFNDMKVRKSST QEEIKKRKKAVLFCLSDDKRQIIVEEAKQILVGDI GDTVEDPYTSFVKLLPLNDCRYALYDATYETKE SKKEDLVFIFWAPESAPLKSKMIYASSKDAIKKK FTGIKHEWQVNGLDDIKDRSTLGEKLGGNVVVS LEGKPL
3176	A	99	1567	PRGCWSSCLDAMFRLNSLSALAELAVGSRWYH GGSQPIQIRRRLMMVAFLGASAVTASTGLLWKR AHAESPPCVDNLKSDIGDKGKNKDEGDVCNHEK KTADLAPHPEEKKKKRSGFRDRKVMEYENRIRA YSTPDKIFRYFATLKVISEPGEAEVFMTPEDFVRS ITPNEKQPEHLGLDQYIIKRFDGKTEKISQEREKF ADEGSIFYTLGECGLISFSDYIFLTTVLSTPQRNFE IAFKMFDLNGDGEVDMEEFEQVQSIIRSQTSMG MRHRDRPTTGNTLKSGLCSALTTYFFGADLKGK LTIKNFLEFQRKLQHDVLKLEFERHDPVDGRITE RQFGGMLLAYSGVQSKKLTAMQRQLKKHFKEG KGLTFQEVENFFTFLKNINDVDTALSFYHMAGAS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LDKVTMQQVARTVAKVELSDHVCDVVFALFDCDGNGELSNKEFVSIMKQRLMRGLEKPKDMGFTRLMQAMWKCAQETAWDFALPKQ
3177	A	182	648	LGVVGSGAAVGGRQAARGAALGRRPMAAVLGALGATRRLLAALRGQSLGLAAMSSGTHRLTAEERNQAILDLKAAGWSELSERDAIYKEFSFHNFNQAFGFMSRVALQAEKMNHHPEWFNVYNKVQITLTSHDCGELTKKDVKLAKFIEKAAASV
3178	A	8	612	ACGCRSFCGSTVMSLLLYYALPALGSYAMLSIFFLRRPHLLHTPRAPTFRIRLGAHRGGSGELLENTMEAMENSMAQRSDLLELDCQLTRDRVVVVSHDENLCRQSGLNRDVGSLDFEDLPLYKEKLEVYFSPGHFAHGSDRRMVRLEDLFQRFPRTPMSVEIKGKNEELIREIAGLVRRYDRNEITIWASEKSSVMKKCK
3179	A	88	1496	QETSKMETLSFPRYNVAEIVIHIRNKILTGADGKNLTKNDLYPNPKPEVLHMIYMRALQIVYGIRLEHFYMMPVNSEVMYPHLMEGFLPFSNLVTHLDSFLPICRVNDFETADILCPKAKRTSRFLSGIINFIHFREACRETYMEFLWQYKSSADKMQQLNAAHQEALMKLERLDSVPVEEQEEFKQLSDGIQELQQSLNQDFHQKTIVLQEGNSQKKSNISEKTKRLNELKLSVVSLKEIQESLKTKIVDSPEKLKNYKEKMKDTVQKLKNARQEVVEKYEIYGDSVDCLPSCQLEVQLYQKKIQDLSDNREKLASILKESLNLEDQIESDESELKKLKTEENSFKRLMIVKKEKLATAQFKINKKHEDVKQYKRTVIEDCNKVQEKRGAVYERVTTINHEIQKIRLGIQQLKDAADREKLKSQEIFLNLKTALEKYHDGIEKAAEDSYAKIDEKTAELKRKMFKMST
3180	A	298	7086	GNMACWPQLRLLLWKNLTFRRRQTCQLLLEVAWPLFIFLILISVRLSYPPYEQHECHFPNKAMPSAGTLPWVQGIICNANNPCFRYPTPGEAPGVVGNFNKSIVARLFSDARRLLLYSQKDTSMKDMRKVLRTLQQIKKSSSNLKLQDFLVDNETFSGFLYHNLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMIQLGDQEVSELCGLPREKLAAAERVLRSNMDILKPILRTLNSTSPFPSKELAEATKTLLHSLGTLAQELFSMRSWSDMRQEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEGGGLKIKSLNWYEDNNYKALFGGNGTEEDAETFYDNSTTPYCNDLMKNLESSPLSRIIWKALKPLLVGKILYTPDTPATRQVMAEVNKTFQELAVFHDLEGMWEELSPKIWTFMENSQEMDLVRMLLDSRDNDHFWEQQLDGLDWTAQDIVAFLAKHPEDVQSSNGSVYTWREAFNETNQAIRTISRFMECVNLNKLEPIATEVWLINKSMELLDERKFWAGIVFTGITPGSIELPHHVKYKIRMGIDNVERTNKIKDGYWDPGPRADPFEDMRYVWGGFAYLQDVVEQAIIRVLTGTEKKTGVYMQQMPYPCYVDDIFLRVMSRSMPLFMTLAWIYSVAVIIKGIVYEKEARLKETMRIMGLDNSILWFSWFISSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFVFLSVFAVVTILQCFLISTLFSRANLAAACGGIIYFTLYLPYVLCVAWQDYVGFTLKIFASLLSPVAFGFGCEYFALFEEQGIGVQWDNLFESPVEEDGFNLTTSVSMMLFDTFLYGVMTWYIEAVFPGQYGIPRPWYFPCTKSYWFGEESDEKSHPGSNQKRISEICMEEEPTHLKLGVSIQNLVKVY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RDGMKVAVDGLALNFYEGQITSFLGHNGAGKTTTMSILTGLFPPTSGTAYILGKDIRSEMSTIRQNLGVCPQHNVLFDMLTVEEHIWFYARLKGLSEKHVKAEMEQMALDVGLPSSKLKSKTSQLSGGMQRKLSVALAFVGGSKVVILDEPTAGVDPYSRRGIWELLLKYRQGRTIILSTHHMDEADVLGDRIAIISHGKLCCVGSSLFLKNQLGTGYYLTLVKKDVESSLSSCRNSSSTVSYLKKEDSVSQSSSDAGLGSDHESDTLTIDVSAISNLIRKHVSEARLVEDIGHELTYVLPYEAAKEGAFVELFHEIDDRLSDLGISSYGISETTLEEIFLKVAEESGVDAETSDGTLPARRNRRAFGDKQSCLRPFTEDDAADPNDSDIDPESRETDLLSGMDGKGSYQVKGWKLTQQQFVALLWKRLLIARRSRKGFFAQIVLPAVFVCIALVFSLIVPPFGKYPSLELQPWMYNEQYTFVSNDAPEDTGTLELLNALTKDPGFGTRCMEGNPIPDTPCQAGEEEWTTAPVPQTIMDLFQNGNWTMQNPSPACQCSSDKIKKMLPVCPPGAGGLPPPQRKQNTADILQDLTGRNISDYLVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVSNTQALPPSQEVNDATKQMKKHLKLAKDSSADRFLNSLGRFMTGLDTRNNVKVWFNNKGWHAISSFLNVINNAILRANLQKGENPSHYGITAFNHPLNLTKQQLSEVAPMTTSVDVLVSICVIFAMSFVPASFVVFLIQERVSKAKHLQFISGVKPVIYWLSNFVWDMCNYVVPATLVIIIFICFQQKSYVSSTNLPVLALLLLLYGWSITPLMYPASFVFKIPSTAYVVLTSVNLFIGINGSVATFVLELFTDNKLNNINDILKSVFLIFPHFCLGRGLIDMVKNQAMADALERFGENRFVSPLSWDLVGRNLFAMAVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLNDEDEDVRRERQRILDGGGQNDILEIKELTKIYRRKRKPAVDRICVGIPPGECFGLLGVNGAGKSSTFKMLTGDTTVTRGDAFLNRNSILSNIHEVHQNMGYCPQFDAITELLTGREHVEFFALLRGVPEKEVGKVGEWAIRKLGLVKYGEKYAGNYSGGNKRKLSTAMALIGGPPVVFLDEPTTGMDPKARRFLWNCALSVVKEGRSVVLTSHSMEECEALCTRMAIMVNGRFRCLGSVQHLKNRFGDGYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPKEKHRNMLQYQLPSSLSSLARIFSILSQSKKRLHIEDYSVSQTTLDQVFVNFAKDQSDDDHLKDLSLHKNQTVVDVAVLTSFLQDEKVKESYV
3181	A	215	1367	PPATSQAALPEALSKGRETPRPATHPARSQDVRPLSCPFDFLRDNVEWSEEQAAAAERKVQENSIQRVCQEKQVDYEINAHKYWNDFYKIHENGFFKDRHWLFTEFPELAPSQNQNHLKDWFLENKSEVPECRNNEDGPGLIMEEQHKCSSKSLEHKTQTPPVEENVTQKISDLEICADEFPGSSATYRILEVGCGVGNTVFPILQTNNDPGLFVYCCDFSSTAIELVQTNSEYDPSRCFAFVHDLCDEEKSYPVPKGSLDIIILIFVLSAIVPDKMQKAINRLSRLLKPGGMVLLRDYGRYDMAQLRFKKGQCLSGNFYVRGDGTRVYFFTQEELDTLFTTAGLEKVQNLVDRRLQVNRGKQLTMYRVWIQCKYCKPLLSSTS
3182	A	3	1289	GSETQHLPRDPQHLPWDPQQHQDRRRPELFHAFARDSAPPPSMVLAAETTSQQERLQAIAEKRKRQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AEIENKRRQLEDERRQLQHLKSKALRERWLLEG TPSSASEGDEDLRRQMQDDEQKTRLLEDSVSRLE KGIEVLERGDSAPAAAKENAAAPSPVRAPAPSPA KEERKTEVVMNSQQTPVGTPKDKRVSNTPLRTV DGSPMMKAAMYSVEITVEKDKVTGETRVLSSTT LLPRQPLPLGIKVYEDETKVVHAVDGTAENGIHP LSSSEVDELIHKADEVTLSEAGSTAGAAETRGAV EGAARTTPSRREITGVQAQPGEATSGPPGIQPGQE PPVTMIFMGYQNVEDEAETKKVLGLQDTITAEL VVIEDAAEPKEPAPPNGSAAEPPTEAASREENQA GPEATTSDPQDLDMKKHRCKCCSIM
3183	A	333	1931	IAPTGGSHSEIQKQLGSGGDSSSQRRAERRTEPRS APRPRWGRSARSPGAHKLPGPPRRRDPGAWARL EAAAAHRHSRGSMGRRMRGAAATAGLWLLAL GSLLALWGGLLPPRTELPASRPPEDRLPRRPARS GGPAPAPRFPLPPPLAWDARGGSLKTFRALLTLA AGADGPPRQSRSEPRWHVSARQPRPEESAAVHG GVFWSRGLEEQVPPGFSEAQAAAWLEAARGAR MVALERGGCGRSSNRLARFADGTRACVRYGINP EQIQGEALSYYLARLLGLQRHVPPLALARVEAR GAQWAQVQEELRAAHWTEGSVVSLTRWLPNLT DVVVPAPWRSEDGRLRPLRDAGGELANLSQAEL VDLVQWTDLILFDYLTANFDRLVSNLFSLQWDP RVMQRATSNLHRGPGGALVFLDNEAGLVHGYR VAGMWDKYNEPLLQSVCVFRERTARRVLELHR GQDAAARLLRLYRRHEPRFPELAALADPHAQLL QRRLDFLAKHILHCKAKYGRRSGDLVSPGGKER DLGLGYG
3184	A	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLVAA ALLVGFILFLTRSRGRAASAGQEPLHNEELAGAG RVAQPGPLEPEEPRAGGRPRRRRDLGSRLQAQR RAQRVAWAEADENEEEAVILAQEEEGVEKPAET HLSGKIGAKKLRKLEEKQARKAQREAEEAEREE RKRLESQREAEWKKEEERLRLEEEQKEEEERKA REEQAQREHEEYLKLKEAFVVEEEGVGETMTEE QSQSFLTEFINYIKQSKVVLLEDLASQVGLRTQD TINRIQDLLAEGTITGVIDDRGKFIYITPEELAAVA NFIRQRGRVSIAELAQASNSLIAWGRESPAQAPA
3185	A	2981	7173	CLLAGKFSSTLYETGGCDMSLVNFEPAARRASNI CDTDSHVSSSTSVRFYPHDVLSLPQIRLNRLLTID TDLLEQQDIDLSPDLAATYGPTEEAAQKVKHYY RFWILPQLWIGINFDRLTLLALFDRNREILENVLA VILAILVAFLGSILLIQGFFRDIWVFQFCLVIASCQ YSLLKSVQPDSSSPRHGHNRIIAYSRPVYFCICCG LIWLLDYGSRNLTATKFKLYGITFTNPLVFISARD LVIVFTLCFPIVFFIGLLPQVNTFVMYLCEQLDIHI FGGNATTSLLAALYSFICSIVAVALLYGLCYGAL KDSWDGQHIPVLFSIFCGLLVAVSYHLSRQSSDP SVLFSLVQSKIFPKTEEKNPEDPLSEVKDPLPEKL RNSVSERLQSDLVVCIVIGVLYFAIHVSTVFTVLQ PALKYVLYTLVGFVGFVTHYVLPQVRKQLPWH CFSHPLLKTLEYNQYEVRNAATMMWFEKLHVW LLFVEKNIIYPLIVLNELSSSAETIASPKKLNTELG ALMITVAGLKLLRSSFSSPTYQYVTVIFTVLFFKF DYEAFSETMLLDLFFMSILFNKLWELLYKLQFVY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TYIAPWQITWGSAFHAFAQPFAVPHSAMLFIQAA VSAFFSTPLNPFLGSAIFITSYVRPVKFWERDYNT KRVDHSNTRLASQLDRNPGTYCQQREVEAITEG VEEDEGFCCCEPGHIPHMLSFNAAFSQRWLAWE VIVTKYILEGYSITDNSAASMLQVFDLRKVLTTY YVKGIIYYVTTSSKLEEWLANETMQEGLRLCAD RNYVDVDPTFNPNIDEDYDHRLAGISRESFCVIY LNWIEYCSSRRAKPVDVDKDSSLVTLCYGLCVL GRRALGTASHHMSSNLESFLYGLHALFKGDFRIS SIRDEWIFADMELLRKVVVPGIRMSIKLHQDHFT SPDEYDDPTVLYEAIVSHEKNLVIAHEGDPAWRS AVLANSPSLLALRHVMDDGTNEYKIIMLNRRYL SFRVIKVNKECVRGLWAGQQQELVFLRNRNPER GSIQNAKQALRNMINSSCDQPIGYPIFVSPLTTSY SDSHEQLKDILGGPISLGNIRNFIVSTWHRLRKGC GAGCNSGGNIEDSDTGGGTSCTGNNATTANNPH SNVTQGSIGNPGQGSGTGLHPPVTSYPPTLGTSHS SHSVQSGLVRQSPARASVASQSSYCYSSRHSSLR MSTTGFVPCRRSSTSQISLRNLPSSIQSRLSMVNQ MEPSGQSGLACVQHGLPSSSSSSQSIPACKHHTL VGFLATEGGQSSATDAQPGNTLSPANNSHSRKA EVIYRVQIVDPSQILEGINLSKRKELQWPDEGIRL KAGRNSWKDWSPQEGMEGHVIHRWVPCSRDPG TRSHIDKAVLLVQIDDKYVTVIETGVLELGAEV
3186	A	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDCGSV DGVIKEVNVSPCPTQPCQLSKGQSYSVNVTFTSN IQSKSSKAVVHGILMGVPVPFPIPEPDGCKSGINC PIQKDKTYSYLNKLPVKSEYPSIKLVVEWQLQDD KNQSLFCWEIPVQIVSHL
3187	A	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDCGSV DGVIKEVNVSPCPTQPCQLSKGQSYSVNVTFTSN IQSKSSKAVVHGILMGVPVPFPIPEPDGCKSGINC PIQKDKTYSYLNKLPVKSEYPSIKLVVEWQLQDD KNQSLFCWEIPVQIVSHL
3188	A	2	3483	PRVRTKLILLVNDKKRYERVGGGPKRLGRDVEM EEMIEQLQEKVHELEKQNDTLKNRLISAKQQLQT QGYRQTPYNNVQSRINTGRRKANENAGLQECPR KGIKFQDADVAETPHPMFTKYGNSLLEEARGEIR NLENVIQSQRGQIEELEHLAEILKTQLRRKENEIE LSLLQLREQQATDQRSNIRDNVEMIKLHKQLVE KSNALSAMEGKFIQLQEKQRTLKISHDALMANG DELNMQLKEQRLKCCSLEKQLHSMKFSERRIEEL QDRINDLEKERELLKENYDKLYDSAFSAAHEEQ WKLKEQQLKVQIAQLETALKSDLTDKTEILDRL KTERDQNEKLVQENRELQLQYLEQKQQLDELKK RIKLYNQENDINADELSEALLLIKAQKEQKNGDL SFLVKVDSEINKDLERSMRELQATHAETVQELEK TRNMLIMQHKINKDYQMEVEAVTRKMENLQQD YELKVEQYVHLLDIRAARIHKLEAQLKDIAYGTK QYKFKPEIMPDDSVDEFDETIHLERGENLFEIHIN KVTFSSEVLQASGDKEPVTFCTYAFYDFELQTTP VVRGLHPEYNFTSQYLVHVNDLFLQYIQKNTITL EVHQAYSTEYETIAACQLKFHEILEKSGRIFCTAS LIGTKGDIPNFGTVEYWFRLRVPMDQAIRLYRER AKALGYITSNFKGPEHMQSLSQQAPKTAQLSSTD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				STDGNLNELHITIRCCNHLQSRASHLQPHPYVVYKFFDFADHDTAIIPSSNDPQFDDHMYFPVPMNMDLDRYLKSESLSFYVFDDSDTQENIYIGKVNVPLISLAHDRCISGIFELTDHQKHPAGTIHVILKWKFAYLPPSGSITTEDLGNFIRSEEPEVVQRLPPASSVSTLVLAPRPKPRQRLTPVDKKVSFVDIMPHQSDVSQEGSVDEVKENTEKMQQGKDDVSLLSEGQLAEQSLASSEDETEITEDLEPEVEEDMSASDSDDCIIPGPISKNIKQPSEKIRIEIIALSLNDSQVTMDDTIQRLFVECRFYSLPAEETPVSLPKPKSGQWVYYNYSNVIYVDKENNKAKRDILKAILQKQEMPNRSLRFTVVSDPPEDEQDLECEDIGVAHVDLADMFQEGRDLIEQNIDVFDARADGEGIGKLRVTVEALHALQSVYKQYRDDLEA
3189	A	476	1175	MKGSGWHLRSGMVGTLITTILPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIYQCQIYRSLLALPQDLQAARALMGISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTLFILAGLLCMGAVSWTTNDVVQNFYNPLLPSGMKFEIGQALYLGFISSSLSLIGGTLLCLSCQDEAPYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV
3190	A	267	1037	DRMAWQGLVLAACLLMFPSTTADCLSRCSLCAVKTQDGPKPINPLICSLQCQAALLPSEEWERCQSFLSFFTPSTLGLNDKEDLGSKSVGEGPYSELAKLSGSFLKELEKSKFLPSISTKENTLSKSLEEKLRGLSDGFREGAESELMRDAQLNDGAMETGTLYLAEEDPKEQVKRYGGFLRKYPKRSSEVAGEGDGDSMGHEDLYKRYGGFLRRIRPKLKWDNQKRYGGFLRRQFKVVTRSQEDPNAYSGELFDA
3191	A	29	574	GTSAGAQTKGALCQLKVPTEKLPSPLPTMADEIDFTTGDAGASSTYPMQCSALRKNGFVVLKGRPCKIVEMSTSKTGKHGHAKVHLVGIDIFTGKKYEDICPSTHNMDVPNIKRNDYQLICIQDGYLSLLTETGEVREDLKLPEGELGKEIEGKYNAGEDVQVSVMCAMSEEYAVAIKPCK
3192	A	105	1661	KVSADGMQSCESSGDSADDPLSRGLRRRGQPRVVVIGAGLAGLAAAKALLEQGFTDVTVLEASSHIGGRVQSVKLGHATFELGATWIHGSHGNPIYHLTEANGLLEETTDGERSVGRISLYSKNGVACYLTNHGRRIPKDVVEEFSDLYNEVYNLTQEFFRHDKPVNAESQNSVGVFTREEVRNRIRNDPDDPEATKRLKLAMIQQYLKVESCESSSHSMDEVSLSAFGEWTEIPGAHHIIPSGFMRVVELLAEGIPAHVIQLGKPVRCIHWDQASARPRGPEIEPRGEGDHNHDTGEGGQGGEEPRGGRWDEDEQWSVVVECEDCELIPADHVIVTVSLGVLKRQYTSFFRPGLPTEKVAAIHRLGIGTTDKIFLEFEEPFWGPECNSLQFVWEDEAESHTLTYPPELWYRKICGFDVLYPPERYGHVLSGWICGEEALVMEKCDDEAVAEICTEMLRQFTGNPNIPKPRRILRSAWGSNPYFRGSYSYTQVGSSGADVEKLAKPLPYTESSKTATK
3193	A	1	1928	QLGTRRCLRGDKVTNAMQDFLVTNLEPRFIEPQTANLSVVFKDSNSTTPLIFVLSPGTDPAADLYKFAEEMKFSKKLSAISLGQGQGPRAEAMMRSSIERGKWVFFQNCHLAPSWMPALERLIEHINPDKVHRDF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RLWLTSLPSNKFPVSILQNGSKMTIEPPRGVRANLLKSYSSLGEDFLNSCHKVMEFKSLLLSLCLFHGNALERRKFGPLGFNIPYEFTDGDLRICISQLKMFLDEYDDIPYKVLKYTAGEINYGGRVTDDWDRRCIMNILEDFYNPDVLSPEHSYSASGIYHQIPPTYDLHGYLSYIKSLPLNDMPEIFGLHDNANITFAQNETFALLGTIIQLQPKSSSAGSQGREEIVEDVTQNILLKVPEPINLQWVMAKYPVLYEESMNTVLVQEVIRYNRLLQVITQTLQDLLKALKGLVVMSSQLELMAASLYNNTVPELWSAKAYPSLKPLSSWVMDLLQRLDFLQAWIQDGIPAVFWISGFFFPQAFLTGTLQNFARKFVISIDTISFDFKVMFEAPSELTQRPQVGCYIHGLFLEGARWDPEAFQLAESQPKELYTEMAVIWLLPTPNRKAQDQDFYLCPIYKTLTRAGTLSTTGHSTNYVIAVEIPTHQPQRHWIKRGVALICALDY
3194	A	1	1023	DGWTPVHAAVDTGNVDSLKLLMYHRIPAHGNSFNEEESESSVFDLDGGEESPEGISKPVVPADLINHANREGWTAAHIAASKGFKNCLEILCRHGGLEPERRDKCNRTVHDVATDDCKHLLENLNALKIPLRISVGEIEPSNYGSDDLECENTICALNIRKQTSWDDFSKAVSQALTNHFQAISSDGWWSLEDVTCNNTTDSNIGLSARSIRSITLGNVPWSVGQSFAQSPWDFMRKNKAEHITVLLSGPQEGCLSSVTYASMIPLQMMQNYLRLVEQYHNVIFHGPEGSLQDYIVHQLALCLKHRQMGWQDSPVEIVEELEVGCWFFPREQLLRTCSLVA
3195	A	1	1809	MAASAQVSVTFEDVAVTFTQEEWGQLDAAQRTLYQEVMLETCGLLMSLGCPLFKPELIYQLDHRQELWMATKDLSQSSYPGDNTKPKTTEPTFSHLALPEEVLLQEQLTQGASKNSQLGQSKDQDGPSEMQEVHLKIGIGPQRGKLLEKMSSERDGLGSDDGVCTKITQKQVSTEGDLYECDSHGPVTDALIREEKNSYKCEECGKVFKKNALLVQHERIHTQVKPYECTECGKTFSKSTHLLQHLIIHTGEKPYKCMECGKAFNRRSHLTRHQRIHSGEKPYKCSECGKAFTHRSTFVLHHRSHTGEKPFVCKECGKAFRDRPGFIRHYIIHTGEKPYECIECIECGKAFNRRSYLTWHQQIHTGVKPFECNECGKAFCESADLIQHYIIHTGEKPYKCMECGKAFNRRSHLKQHQRIHTGEKPYECSECGKAFTHCSTFVLHKRTHTGEKPYECKECGKAFSDRADLIRHFSIHTGEKPYECVECGKAFNRSSHLTRHQQIHTGEKPYECIQCGKAFCRSANLIRHSIIHTGEKPYECSECGKAFNRGSSLTHHQRIHTGRNPTIVTDVGRPFMTAQTSVNIQELLLGKEFLNITTEENLW
3196	A	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARGTGALLLRGSLLASGRAPRRASSGLPRNTVVLFVPQQEAWVVERMGRFHRILEPGLNILIPVLDRIRYVQSLKEIVINVPEQSAVTLDNVTLQIDGVLYLRIMDPYKASYGVEDPEYAVTQLAQTTMRSELGKLSLDKVFRERESLNASIVDAINQAADCWGIRCLRYEIKDIHVPPRVKESMQMQVEAERRKRATVLESEGTRESAINVAEGKKQAQILASEAEKAEQINQAAGEASAVLAKAKAKAEAIRILAAALTQHNGDAAASLTVAEQYVSAFSKLAKDSNTILLPSNPGDVTSMVAQAMGVYGALTKAPVPGTPDSLSSGSSRDVQGTDASLDE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ELDRVKMS
3197	A	66	3632	LWECAAAAAGQRDGGVTLFLKGRVLGRRCAASLFAREVCVSTSSSRPACFLHCARARGEQMHQMASGVGSMKRSPRKMWRPGEKKEPQGVVYEDVRDDTEDFKEPLKVVFEGSAYGLQNFNKQKKLKTCDDMDTFFLHYAAAEGQIELMEKITRDSSLEVLHEMDDYGNTPLHCAVEKNQIESVKFLLSRGANPNLRNFNMMAPLHIAVQGMNNEVMKVLLEHRTIDVNLEGENGNTAVIIACTTNNSEALQILLNKGAKPCKSNKWGCFPIHQAAFSGSKECMEIILRFGEEHGYSRQLHINFMNNGKATPLHLAVQNGDLEMIKMCLDNGAQIDPVEKGRCTAIHFAATQGATEIVKLMISSYSGSVDIVNTTDGCHETMLHRASLFDHHELADYLISVGADINKIDSEGRSPLILATASASWNIVNLLLSKGAQVDIKDNFGRNFLHLTVQQPYGLKNLRPEFMQMQQIKELVMDEDNDGCTPLHYACRQGGPGSVNNLLGFNVSIHSKSKDKKSPLHFAASYGRINTCQRLLQDISDTRLLNEGDLHGMTPLHLAAKNGHDKVVQLLLKKGALFLSDHNGWTALHHASMGGYTQTMKVILDTNLKCTDRLDEDGNTALHFAAREGHAKAVALLLSHNADIVLNKQQASFLHLALHNKRKEVVLTIIRSKRWDECLKIFSHNSPGNKCPITEMIEYLPECMKVLLDFCMLHSTEDKSCRDYYIEYNFKYLQCPLEFTKKTPTQDVIYEPLTALNAMVQNNRIELLNHPVCKEYLLMKWLAYGFRAHMMNLGSYCLGLIPMTILVVNIKPGMAFNSTGIINETSDHSEILDTTNSYLIKTCMILVFLSSIFGYCKEAGQIFQQKRNYFMDISNVLEWIIYTTGIIFVLPLFVEIPAHLQWQCGAIAVYFYWMNFLLYLQRFENCGIFIVMLEVILKTLLRSTVVFIFLLLAFGLSFYILLNLQDPFSSPLLSIIQTFSMMLGDINYRESFLEPYLRNELAHPVLSFAQLVSFTIFVPIVLMNLLIGLAVGDIAEVQKHASLKRIAMQVELHTSLEKKLPLWFLRKVDQKSTIVYPNKPRSGGMLFHIFCFLFCTGEIRQEIPNADKSLEMEILKQKYRLKDLTFLLEKQHELIKLIIQKMEIISETEDDDSHCSFQDRFKKEQMEQRNSRWNTVLRAVKAKTHHLEP
3198	A	51	2177	KEKSLHHVDQRPPLWHPGRPGTSQSAAMNASSEGESFAGSVQIPGGTTVLVELTPDIHICGICKQQFNNLDAFVAHKQSGCQLTGTSAAAPSTVQFVSEETVPATQTQTTTRTITSETQTITVSAPEFVFEHGYQTYLPTESNENQTATVISLPAKSRTKKPTTPPAQKRLNCCYPGCQFKTAYGMKDMERHLKIHTGDKPHKCEVCGKCFSRKDKLKTHMRCHTGVKPYKCKTCDYAAADSSSLNKHLRIHSDERPFKCQICPYASRNSSQLTVHLRSHTGDAPFQCWLCSAKFKISSDLKRHMRVHSGEKPFKCEFCNVRCTMKGNLKSHIRIKHSGNNFKCPHCAFLGDSKATLRKHSRVHQSEHREKCSECSYSCSSKAALRIHERIHCTVRPFKCNYCSFDSKQPSNLSKHMKKFHGDMVKTEALERKDTGRQSSRQVAKLDAKKSFHCDICDASFMREDSLRSHKRQHSEYNESKNSDVTVLQFQIDPSKQPATPLTVGHLQVPLQPSQVPQFSEGRVKIIVGHQVPQANTIVQAAAAAVNIVPPALVAQNPEELPGNSRLQILRQVSLIAPPQSSRCPSEAGAMTQPAVLLTTHEQTD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GATLHQTLIPTASGGPQEGSGNQTFITSSGITCTD FEGLNALIQEGTAEVTVVSDGGQNIAVATTAPPV FSSSSQQELPKQTYSIIQGAAHPALLCPADSIPD
3199	A	13	2247	QSFHSMEGDPSGLPLLARGASCYSLICPCPRPAD WSILQGTDWSILQSADWCIYNPLARHRALTGVFL QSADWCTYNPLARQKSSPSPHSTQEVQLASPLTR RPNKKDSAERNHRPAREGSVAQRQPNPAALEKA EPAARKRNEREGGGSQEPGREHSLEKGYWAPGL GPDPSMCSKQVDPSEGASSHLKHRGGSRAAHLE VRRLLRRLVGALVAEAGFCYVQVAEGQRVVGV LEVAEAAAAPVQHEPTAAVATQSRWFPRGTRPG LCSLPIAVAALLCPGSGPGAQSGLEFVERPPPSPL AVVLARWPLPPPAGRCPRDAPEARVPEKARAEG SERENNYGCGVVGGEMTTLVLDNGAYNAKIGY SHENVSVIPNCQFRSKTARLKTFTANQIDEIKDPS GLFYILPFQKGYLVNWDVQRQVWDYLFGKEMY QVDFLDTNIIITEPYFNFTSIQESMNEILFEEYQFQ AVLRVNAGALSAHRYFRDNPSELCCIIVDSGYSF THIVPYCRSKKKKEAIIRINVGGKLLTNHLKEIISY RQLHVMDETHVINQVKEDVCYVSQDFYRDMDI AKLKGEENTVMIDYVLPDFSTIKKGFCKPREEMV LSGKYKSGEQILRLANERFAVPEILFNPSDIGIQE MGIPEAIVYSIQNLPEEMQPHFFKNIVLTGGNSLF PGFRDRVYSEVRCLTPTDYDVSVVLPENPITYAW EGGKLISENDDFEDMVVTREDYEENGHSVCEEK FDI
3200	A	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIWKTR SCAGISGKSQLLFALVFTTRYLDLFTSFISLYNTS MKVWYAIHRNVFHLQCTGLWTLNLCQLCIFN
3201	A	1	469	IRHEGRGQRGKMELVQVLKRGLQQITGHGGLRG YLRVFFRTNDAKVGTLVGEDKYGNKYYEDNKQ FFGRHRWVVYTTEMNGKNTFWDVDGSMVPPE WHRWLHSMTDDPPTTKPLTARKFIWTNHKFNVT GTPEQYVPYSTTRKKIQEWIPPSTPYK
3202	A	144	840	NSSQRIMATHALEIAGLFLGGVGMVGTVAVTVM PQWRVSAFIENNIVVFENFWEGLWMNCVRQANI RMQCKIYDSLLALSPDLQAARGLMCAASVMSFL AFMMAILGMKCTRCTGDNEKVKAHILLTAGIIFII TGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELG EALYLGWTTALVLIVGGALFCCVFCCNEKSSSYR YSIPSHRTTQKSYHTGKKSPSVYSRSQYV
3203	A	2	473	KYRYRRPYPVMRKICQVGPAGLAFILNISPVAHR VALCHLAGCQEQAAWYHTLQILFFLVSAYFFSCP VPEKYFPGSCDIVGHGHQIFHAFLSICTLSQLEAIL LDYQGRQEIFLQRHGPLSVHMACLSFFFLAACSA ATAALLRHKVKARLTKKDS
3204	A	1808	668	PESAPLPAFISSRILPAAWRNWCSYVVTRTISCHV QNGTYLQRVLQNCPWPMSCPGSSYRTVVRPTYK VMYKIVTAREWRCCPGHSRVSCEEVAGSSASLE PMWSGSTMRRMALRPTAFSGCLNCSKVSELTER LKVLEAKMTMLTVIEQPVPPTPATPEDPAPLWGP PPAQGSPGDGGLQDQVGAWGLPGPTGPKGDAG SRGPMGMRGPPGDPLLSNTFTETNNHWPQGPTG PPGPPGPMGPPGPPGPTGVPGSPGHIGPPGPTGPK GISGHPGEKGERGLRGEPGPQGSAGQRGEPGPKG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DPGEKSHWGEGLHQLREALKILAERVLILETMIG LYEPELGSGAGPAGTGTPSLLRGKRGGHATNYRI VAPRSRDERG
3205	A	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRMRM KYGGQEFWADLNAMNVYETTEFDQLRRLSTPPS SNVNSIYHTVWKFFCRDHFGWREYPESVIRLIEE ANSRGLKEVRFMMWNNHYILHNSFFRREIKRRP LFRSCFILLPYLQTLGGVPTQAPPPLEATSSSQIICP DGVTSANFYPETWVYMHPSQDFIQVPVSAEDKS YRIIYNLFHKTVPEFKYRILQILRVQNQFLWEKY KRKKEYMNRKMFGRDRIINERHLFHGTSQDVVD GICKHNFDPRVCGKHATMFGQGSYFAKKASYSH NFSKKSSKGVHFMFLAKVLTGRYTMGSHGMRR PPPVNPGSVTSDLYDSCVDNFFEPQIFVIFNDDQS YPYFVIQYEEVSNTVSI
3206	A	297	4500	CLVDSKLWKGARSVYHQLFMSSLLMDLKYKKL FAVRFAKNYERLQSDYVTDDHDREFSVADLSVQ IFTVPSLARMLITEENLMSIIIKTFMDHLRHRDAQ GRFQFERYTALQAFKFRRVQSLILDLKYVLISKPT EWSDELRQKFLEGFDAFLELLKCMQGMDPITRQ VGQHIEMEPEWEAAFTLQMKLTHVISMMQDWC ASDEKVLIEAYKKCLAVLMQCHGGYTDGEQPIT LSICGHSVETIRYCVSQEKVSIHLPVSRLLAGLHV LLSKSEVAYKFPELLPLSELSPPMLIEHPLRCLVL CAQVHAGMWRRNGFSLVNQIYYYHNVKCRRE MFDKDVVMLQTGVSMMDPNHFLMIMLSRFELY QIFSTPDYGKRFSSEITHKDVVQQNNTLIEEMLYL IIMLVGERFSPGVGQVNATDEIKREIIHQLSIKPM AHSELVKSLPEDENKETGMESVIEAVAHFKKPGL TGRGMYELKPECAKEFNLYFYHFSRAEQSKAEE AQRKLKRQNREDTALPPPVLPPFCPLFASLVNILQ SDVMLCIMGTILQWAVEHNGYAWSESMLQRVL HLIGMALQEEKQHLENVTEEHVVTFTFTQKISKP GEAPKNSPSILAMLETLQNAPYLEVHKDMIRWIL KTFNAVKKMRESSPTSPVAETEGTIMEESSRDKD KAERKRKAEIARLRREKIMAQMSEMQRHFIDEN KELFQQTLELDASTSAVLDHSPVASDMTLTALGP AQTQVPEQRQFVTCILCQEEQEVKVESRAMVLA AFVQRSTVLSKNRSKFIQDPEKYDPLFMHPDLSC GTHTSSCGHIMHAHCWQRYFDSVQAKEQRRQQ RLRLHTSYDVENGEFLCPLCECLSNTVIPLLLPPR NIFNNRLNFSDQPNLTQWIRTISQQIKALQFLRKE ESTPNNASTKNSENVDELQLPEGFRPDFRPKIPYS ESIKEMLTTFGTATYKVGLKVHPNEEDPRVPIMC WGSCAYTIQSIERILSDEDKPLFGPLPCRLDDCLR SLTRFAAAHWTVASVSVVQGHFCKPFASLVPND SHEELPCILDIDMFHLLVGLVLAFPALQCQDFSGI SLGTGDLHIFHLVTMAHIIQILLTSCTEENGMDQE NPPCEEESAVLALYKTLHQYTGSALKEIPSGWHL WRSVRAGIMPFLKCSALFFHYLNGVPSPPDIQVP GTSHFEHLCSYLSLPNNLICLFQENSEIMNSLIES WCRNSEVKRYLEGERDAIRYPRESNKLINLPEDY SSLINQASNFSCPKSGGDKSRAPTLCLVCGSLLCS QSYCCQTELEGEDVGACTAHTYSCGSGVGIFLR VRECQVLFLAGKTKGCFYSPPYLDDYGETDQGL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RRGNPLHLCKERFKKIQKLWHQHSVTEEIGHAQ EANQTLVGIDWQHL
3207	A	49	963	QLSPSQAPAGAQEVARRVTVGSASHGGRRSTMA TTVSTQRGPVYIGELPQDFLRITPTQQQRQVQLD AQAAQQLQYGGAVGTVGRLNITVVQAKLAKNY GMTRMDPYCRLRLGYAVYETPTAHNGAKNPRW NKVIHCTVPPGVDSFYLEIFDERAFSMDDRIAWT HITIPESLRQGKVEDKWYSLSGRQGDDKEGMINL VMSYALLPAAMVMPPQPVVLMPTVYQQGVGY VPITGMPAVCSPGMVPVALPPAAVNAQPRCSEE DLKAIQDMFPNMDQEVIRSVLEAQRGNKDAAIN SLLQMGEEP
3208	A	54	1196	LERTPASADMAWTKYQLFLAGLMLVTGSINTLS AKWADNFMAEGCGGSKEHSFQHPFLQAVGMFL GEFSCLAAFYLLRCRAAGQSDSSVDPQQPFNPLL FLPPALCDMTGTSLMYVALNMTSASSFQMLRGA VIIFTGLFSVAFLGRRLVLSQWLGILATIAGLVVV GLADLLSKHDSQHKLSEVITGDLLIIMAQIIVAIQ MVLEEKFVYKHNVHPLRAVGTEGLFGFVILSLLL VPMYYIPAGSFSGNPRGTLEDALDAFCQVGQQP LIAVALLGNISSIAFFNFAGISVTKELSATTRMVL DSLRTVVIWALSLALGWEAFHALQILGFLILLIGT ALYNGLHRPLLGRLSRGRPLAEESEQERLLGGTR TPINDAS
3209	A	104	1999	AKVVSLKEFSCFWRREKPVSSLSSLQVKAEASW DSAVHGCPQLSRGTPVDERLFLIVRVTVQLSHPA DMQLVLRKRICVNVHGRQGFAQSLLKKMSHRSS IPGCGVTFEIVSNIPEDAQGVEEREALARMAANV ENPASADSEAYIEKYLRSVLAVENLLTLDRLRQE VAVKEQLTGKGKLSRRSISSPNVNRLSGSRQDLIP SYSLGSNKGRWESQQDVSQTTVSRGIAPAPALSV SPQNNHSPDPGLSNLAASYLNPVKSFVPQMPKLL KSLFPVRDEKRGKRPSPLAHQPVPRIMVQSASPDI RVTRMEEAQPEMGPDVLVQTMGAPALKICDKP AKVPSPPPVIAVTAVTPAPEAQDGPPSPLSEASSG YFSHSVSTATLSDALGPGLDAAAPPGSMPTAPEA EPEAPISHPPPPTAVPAEEPPGPQQLVSPGRERPDL EAPAPGSPFRVRRVRASELRSFSRMLAGDPGCSP GAEGNAPAPGAGGQALASDSEEADEVPEWLREG EFVTVGAHKTGVVRYVGPADFQEGTWVGVELD LPSGKNDGSIGGKQYFRCNPGYGLLVRPSRVRR ATGPVRRRSTGLRLGAPEARRSATLSGSATNLAS LTAALAKADRSHKNPENRKSWAS
3210	A	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALVVS GGIVGYVKTGSVPSLAAGLLFGSLAGLGAYQLY QDPRNVWGFLAATSVTFVGVMGMRSYYYGKF MPVGLIAGASLLMAAKVGVRMLMTSD
3211	A	1078	594	VGMELPAVNLKVILLGHWLLTTWGCIVFSGSYA WANFTILALGVWAVAQRDSIDAISMFLGGLLATI FLDIVHISIFYPRVSLTDTGRFGVGMAILSLLLKPL SCCFVYHMYRERGGELLVHTGFLGSSQDRSAYQ TIDSAEAPADPFAVPEGRSQDARGY
3212	A	1	1962	FRCGLAPKGRPRRRADPVASAIMDPAEAVLQEK ALKFMMEFRSWCPGWNTMARSRLTATSTSRVQ CSMPRSLWLGCSSLADSMPSLRCLYNPGTGALT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AFQNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCA RLCLNQETVCLASTAMKTENCVAKTKLANGTSS MIVPKQRKLSASYEKEKELCVKYFEQWSESDQV EFVEHLISQMCHYQHGHINSYLKPMLQRDFITAL PARGLDHIAENILSYLDAKSLCAAELVCKEWYR VTSDGMLWKKLIERMVRTDSLWRGLAERRGWG QYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIES NWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIV SGLRDNTIKIWDKNTLECKRILTGHTGSVLCLQY DERVIITGSSDSTVRVWDVNTGEMLNTLIHHCEA VLHLRFNNGMMVTCSKDRSIAVWDMASPTDITL RRVLVGHRAAVNVVDFDDKYIVSASGDRTIKV WNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGS SDNTIRLWDIECGACLRVLEGHEELVRCIRFDNK RIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRT LVEHSGRVFRLQFDEFQIVSSSHDDTILIWDFLND PAAQSEPPRSPSRTYTYISR
3213	A	1	1962	FRCGLAPKGRPRRRADPVASAIMDPAEAVLQEK ALKFMMEFRSWCPGWNTMARSRLTATSTSRVQ CSMPRSLWLGCSSLADSMPSLRCLYNPGTGALT AFQNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCA RLCLNQETVCLASTAMKTENCVAKTKLANGTSS MIVPKQRKLSASYEKEKELCVKYFEQWSESDQV EFVEHLISQMCHYQHGHINSYLKPMLQRDFITAL PARGLDHIAENILSYLDAKSLCAAELVCKEWYR VTSDGMLWKKLIERMVRTDSLWRGLAERRGWG QYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIES NWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIV SGLRDNTIKIWDKNTLECKRILTGHTGSVLCLQY DERVIITGSSDSTVRVWDVNTGEMLNTLIHHCEA VLHLRFNNGMMVTCSKDRSIAVWDMASPTDITL RRVLVGHRAAVNVVDFDDKYIVSASGDRTIKV WNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGS SDNTIRLWDIECGACLRVLEGHEELVRCIRFDNK RIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRT LVEHSGRVFRLQFDEFQIVSSSHDDTILIWDFLND PAAQSEPPRSPSRTYTYISR
3214	A	1	1962	FRCGLAPKGRPRRRADPVASAIMDPAEAVLQEK ALKFMMEFRSWCPGWNTMARSRLTATSTSRVQ CSMPRSLWLGCSSLADSMPSLRCLYNPGTGALT AFQNSSEREDCNNGEPPRKIIPEKNSLRQTYNSCA RLCLNQETVCLASTAMKTENCVAKTKLANGTSS MIVPKQRKLSASYEKEKELCVKYFEQWSESDQV EFVEHLISQMCHYQHGHINSYLKPMLQRDFITAL PARGLDHIAENILSYLDAKSLCAAELVCKEWYR VTSDGMLWKKLIERMVRTDSLWRGLAERRGWG QYLFKNKPPDGNAPPNSFYRALYPKIIQDIETIES NWRCGRHSLQRIHCRSETSKGVYCLQYDDQKIV SGLRDNTIKIWDKNTLECKRILTGHTGSVLCLQY DERVIITGSSDSTVRVWDVNTGEMLNTLIHHCEA VLHLRFNNGMMVTCSKDRSIAVWDMASPTDITL RRVLVGHRAAVNVVDFDDKYIVSASGDRTIKV WNTSTCEFVRTLNGHKRGIACLQYRDRLVVSGS SDNTIRLWDIECGACLRVLEGHEELVRCIRFDNK RIVSGAYDGKIKVWDLVAALDPRAPAGTLCLRT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LVEHSGRVFRLQFDEFQIVSSSHDDTILIWDFLND PAAQSEPPRSPSRTYTYISR
3215	A	2	1376	EARLVGCQRGGPARPGSYSSGAETAGRAMAAN LSRNGPALQEAYVRVVTEKSPTDWALFTYEGNS NDIRVAGTGEGGLEEMVEELNSGKVMYAFCRV KDPNSGLPKFVLINWTGEGVNDVRKGACASHVS TMASFLKGAHVTINARAEEDVEPECIMEKVAKA SGANYSFHKESGRFQDVGPQAPVGSVYQKTNAV SEIKRVGKDSFWAKAEKEEENRRLEEKRRAEEA QRQLEQERRERELREAARREQRYQEQGGEASPQ RTWEQQQEVVSRNRNEQESAVHPREIFKQKERA MSTTSISSPQPGKLRSPFLQKQLTQPETHFGREPA AAISRPRADLPAEEPAPSTPPCLVQAEEEAVYEEP PEQETFYEQPPLVQQQGAGSEHIDHHIQGQGLSG QGLCARALYDYQAADDTEISFDPENLITGIEVIDE GWWRGYGPDGHFGMFPANYVELIE
3216	A	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADLSW DPMAFFTGLWGPFTCVSRVLSHHCFSTTGSLSAI QKMTRVRVVDNSALGNSPYHRAPRCIHVYKKN GVGKVGDQILLAIKGQKKKALIVGHCMPGPRMT PRFDSNNVVLIEDNGNPVGTRIKTPIPTSLRKREG EYSKVLAIAQNFV
3217	A	1	1563	MLCALLLLPSLLGATRASPTSGPQECAKGSTVW CQDLQTAARCGAVGYCQGAVWNKPTAKSLPCD VCQDIAAAAGNGLNPDATESDILALVMKTCEWL PSQESSAGCKWMVDAHSSAILSMLRGAPDSAPA QVCTALSLCEPLQRHLATLRPLSKEDTFEAVAPF MANGPLTFHPRQAPEGALCQDCVRQVSRLQEAV RSNLTLADLNIQEQCESLGPGLAVLCKNYLFQFF VPADQALRLLPPQELCRKGGFCEELGAPARLTQ VVAMDGVPSLELGLPRKQSEMQMKAGVTCEVC MNVVQKLDHWLMSNSSELMITHALERVCSVMP ASITKECIILVDTYSPSLVQLVAKITPEKVCKFIRL CGNRRRARAVHDAYAIVPSPEWDAENQGSFCNG CKRLLTVSSHNLESKSTKRDILVAFKGGCSILPLP YMIQCKHFVTQYEPVLIESLKDMMDPVAVCKKV GACHGPRTPLLGTDQCALGPSFWCRSQEAAKLC NAVQHCQKHVWKEMHLHAGEHA
3218	A	1	1563	MLCALLLLPSLLGATRASPTSGPQECAKGSTVW CQDLQTAARCGAVGYCQGAVWNKPTAKSLPCD VCQDIAAAAGNGLNPDATESDILALVMKTCEWL PSQESSAGCKWMVDAHSSAILSMLRGAPDSAPA QVCTALSLCEPLQRHLATLRPLSKEDTFEAVAPF MANGPLTFHPRQAPEGALCQDCVRQVSRLQEAV RSNLTLADLNIQEQCESLGPGLAVLCKNYLFQFF VPADQALRLLPPQELCRKGGFCEELGAPARLTQ VVAMDGVPSLELGLPRKQSEMQMKAGVTCEVC MNVVQKLDHWLMSNSSELMITHALERVCSVMP ASITKECIILVDTYSPSLVQLVAKITPEKVCKFIRL CGNRRRARAVHDAYAIVPSPEWDAENQGSFCNG CKRLLTVSSHNLESKSTKRDILVAFKGGCSILPLP YMIQCKHFVTQYEPVLIESLKDMMDPVAVCKKV GACHGPRTPLLGTDQCALGPSFWCRSQEAAKLC NAVQHCQKHVWKEMHLHAGEHA
3219	A	1623	572	TSAEGWKGCTCTFKDRSKLREHLRSHTQEKVVA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				CPTCGGMFANNTKFLDHIRRQTSLDQQHFQCSH CSKRFATERLLRDHMRNHVNHYKCPLCDMTCPL PSSLRNHMRFRHSEDRPFKCDCCDYSCKNLIDLQ KHLDTHSEEPAYRCDFENCTFSARSLCSIKSHYR KVHEGDSEPRYKCHVCDKCFTRGNNLTVHLRK KHQFKWPSGHPRFRYKEHEDGYMRLQLVRYES VELTQQLLRQPQEGSGLGTSLNESSLQGIILETVP GEPGRKEEEEEGKGSEGTALSASQDNPSSVIHVV NQTNAQGQQEIVYYVLSEAPGEPPPVPEPPSGGI MEKLQGIAEEPEIQMV
3220	A	2760	745	SLGIPSGNTRGTGLVLDGDTSYTYHLVCMGPEAS GWGQDEPQTWPTDHRAQQGVQRQGVSYSVHA YTGQPSPRGLHSENREDEGWQVYRLGARDAHQ GRPTWALRPEDGEDKEMKTYRLDAGDADPRRL CDLERERWAVIQGQAVRKSSTVATLQGTPDHGD PRTPGPPRSTPLEENVVDREQIDFLAARQQFLSLE QANKGAPHSSPARGTPAGTTPGASQAPKAFNKP HLANGHVVPIKPQVKGVVREENKVRAVPTWAS VQVVDDPGSLASVESPGTPKETPIERBIRLAQERE ADLREQRGLRQATDHQELVEIPTRPLLTKLSLITA PRRERGRPSLYVQRDIVQETQREEDHRREGLHV GRASTPDWVSEGPQPGLRRALSSDSILSPAPDAR AADPAPEVRKVNRIPPDAYQPYLSPGTPQLEFSA FGAFGKPSSLSTAEAKAATSPKATMSPRHLSESS GKPLSTKQEASKPPRGCPQANRGVVRWEYFRLR PLRFRAPDEPQQAQVPHVWGWEVAGAPALRLQ KSQSSDLLERERESVLRREQEVAEERRNALFPEV FSPTPDENSDQNSRSSSQASGITGSYSVSESPFFSPI HLHSNVAWTVEDPVDSAPPGQRKKEQWYAGIN PSDGINSEVLEAIRVTRHKNAMAERWESRIYASE EDD
3221	A	15	478	SRVFFFFFFFPAFKMSKRGRGGSSGAKFRISLGLP VGAVINCADNTGAKNLYIISVKGIKGRLNRLPAA GVGDMVMATVKKGKPELRKKVHPAVVIRQRKS YRRKDGVFLYFEDNAGVIVNNKGEMKGSAITGP VAKECADLWPRIASNAGSIA
3222	A	207	1321	PLIPLHPANRSPATMAELQEVQITEEKPLLPGQTP EAAKTHSVETPYGSVTFTVYGTPKPKRPAILTYH DVGLNYKSCFQPLFQFEDMQEIIQNFVRVHVDAP GMEEGAPVFPLGYQYPSLDQLADMIPCVLQYLN FSTIIGVGVGAGAYILARYALNHPDTVEGLVLINI DPNAKGWMDWAAHKLTGLTSSIPEMILGHLFSQ EELSGNSELIQKYRNIITHAPNLDNIELYWNSYNN RRDLNFERGGDITLRCPVMLVVGDQAPHEDAVV ECNSKLDPTQTSFLKMADSGGQPQLTQPGKLTE AFKYFLQGMGYMASSCMTRLSRSRTASLTSAAS VDGNRSRSRTLSQSSESGTLSSGPPGHTMEVSC
3223	A	132	1664	SARRWGAAGAGPHGLHLRAHGPRPSVRTGLPSV GRQAAGAAMGRGWGFLFGLLGAVWLLSSGHGE EQPPETAAQRCFCQVSGYLDDCTCDVETIDRFNN YRLFPRLQKLLESDYFRYYKVNLKRPCPFWNDIS QCGRRDCAVKPCQSDEVPDGIKSASYKYSEEAN NLIEECEQAERLGAVDESLSEETQKAVLQWTKH DDSSDNFCEADDIQSPEAEYVDLLLNPERYTGYK GPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TSEENTFYSWLEGLCVEKRAFYRLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRRLKNLYFLYLIELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAHKLKEDFRLHFRNISRIMDCVGCFKCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFHLTRQEIVSLFNAFGRISYKCERIRKTSRNLLQNIH
3224	A	2	803	PGSTISWDRDAAGESGTRAASPSPSGSRTAGRLPSPSYSPLPAPSLFPPPPLPAPAASTMSAGGDFGNPLRKFKLVFLGEQSVGKTSLITRFMYDSFDNTYQATIGIDFLSKTMYLEDRTVRLQLWDTAGQERFRSLIPSYIRDSTVAVVVYDITNLNSFQQTSKWIDDVRTERGSDVIIMLVGNKTDLADKRQITIEEGEQRAKELSVMFIETSAKTGYNVKQLFRRVASALPGMENVQEKSKEGMIDIKLDKPQEPPASEGGCSC
3225	A	3	5054	PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKRVAVPNGQPPSAARYMPREVPPRFRCQQDHKVLLKRGQPPPPSCMLLGGGAGPPPCTAPGANPNNAQVTGALLQSESGTAPDSTLGGAAASNYANSTWGSGASSNNGTSPNPIHIWDKVIVDGSDMEEWPCIASKDTESSSENTTDNNSASNPGSEKSTLPGSTTSNKGKGSQCQSASSGNECNLGVWKSDPKAKSVQSSNSTTENNNGLGNWRNVSGQDRIGPGSGFSNFNPNSNPSAWPALVQEGTSRKGALETDNSNSSAQVSTVGQTSREQQSKMENAGVNFVVSGREQAQIHNTDGPKNGNTNSLNLSSPNPMENKGMPFGMGLGNTSRSTDAPSQSTGDRKTGSVGSWGAARGPSGTDTVSGQSNSGNNGNNGKEREDSWKGASVQKSTGSKNDSWDNNNRSTGGSWNFGPQDSNDNKWGEGNKMTSGVSQGEWKQPTGSDELKIGEWSGPNQPNSSTGAWDNQKGHPLLENQGNAQAPCWGRSSSSTGSEVEGQSTGSNHKAGSSDSHNSGRRSYRPTHPDCQAVLQTLLSRTDLDPRVLSNTGWGQTQIKQDTVWDIEEVPRPEGKSDKGTEGWESAATQTKNSGGWGDAPSQSNQMKSGWGELSASTEWKDPKNTGGWNDYKNNNSSNWGGGRPDEKTPSSWNENPSKDQGWGGGRQPNQGWSSGKNGWGEEVDQTKNSNWESSASKPVSGWGEGGQNEIGTWGNGGNASLASKGGWEDCKRSPAWNETGRQPNSWNKQHQQQQPPQQPPPPQPEASGSWGGPPPPPPGNVRPSNSSWSSGPQPATPKDEEPSGWEEPSPQSISRKMDIDDGTSAWGDPNSYNYKNVNLWDKNSQGGPAPREPNLPTPMTSKSASDSKSMQDGWGESDGPVTGARHPSWEEEEDGGVWNTTGSQGSASSHNSASWGQGGKKQMKCSLKGGNNDSWMNPLAKQFSNMGLLSQTEDNPSSKMDLSVGSLSDKKFDVDKRAMNLGDFNDIMRKDRSGFRPPNSKDMGTTDSGPYFEKGGSHGLFGNSTAQSRGLHTPVQPLNSSPSLRAQVPPQFISPQVSASMLKQFPNSGLSPGLFNVGPQLSPQQIAMLSQLPQIPQFQLACQLLLQQQQQQQLLQNQRKISQAVRQQQEQQLARMVSALQQQQQQQRQPGMKHSPSHPVGPKPHLDNMVPNALNVGLPDLQTKGPIPGYGSGFSSGGMDYGMVGGKEAGTESRFKQWTSMMEGLPSVATQEANMHKNGAIVAPGK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TRGGSPYNQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNKIGSKSSNASWPPEFQPGVPWKGIQNIDPESDPYVTPGSVLGGTATSPIVDTDHQLLRDNTTGSNSSLNTSLPSPGAWPYSASDNSFTNVHSTSAKFPDYKSTWSPDPIGHNPTHLSNKMWKNHISSRNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRGWGTQDSRLASASTWSDGGSVRPSYWLVLHNLTPQIDGSTLRTICMQHGPLLTFHLNLTQGTALIRYSTKQEAAKAQTALHMCVLGNTTILAEFATDDEVSRFLAQAQPPTPAATPSAPAAGWQSLETGQNQSDPVGPALNLFGGSTGLGQWSSSAGGSSGADLAGASLWGPPNYSSSLWGVPTVEDPHRMGSPAPLLPGDLLGGGSDSI
3226	A	200	1387	VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTFLPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDSPQDCGFKDHQPLTLQALTVELARWTLMLLLSTAMYGAHAPLLALCHVDGRVPFRPSSAVLLTELTKLLLCAFSLLVGWQAWPQGPPPWRQAAPFALSALLYGANNNLVIYLQRYMDPSTYQVLSNLKIGSTAVLYCLCLRHRLSVRQGLALLLLMAAGACYAAGGLQVPGNTLPSPPPAAAASPMPLHITPLGLLLLILYCLISGLSSVYTELLMKRQRLPLALQNLFLYTFGVLLNLGLHAGGGSGPGLLEGFSGWAALVVLSQALNGLLMSAVMKHGSSITRLFVVSCSLVVNAVLSAVLLRLQLTAAFFLATLLIGLAMRLYYGSR
3227	A	1	679	RSTRARTRRPGLRAVPLPVGGFLGKMKWVWALLLLAALGSGRAERDCRVSSFRVKENFDKARFSGTWYAMAKKDPEGLFLQDNIVAEFSVDETGQMSATAKGRVRLLNNWDVCADMVGTFTDTEDPAKFKMKYWGVASFLQKGNDDHWIVDTDYDTYAVQYSCRLLNLDGTCADSYSFVFSRDPNGLPPEAQKIVRQRQEELCLARQYRLIVHNGYCDGRSERNLL
3228	A	430	1104	QQESPAAGAARMNCKEGTDSSCGCRGNDEKKMLKCVVVGDGAVGKTCLLMSYANDAFPEEYVPTVFDHYAVTVTVGGKQHLLGLYDTAGQEDYNQLRPLSYPNTDVFLICFSVVNPASYHNVQEEWVPELKDCMPHVPYVLIGTQIDLRDDPKTLARLLYMKEKPLTYEHGVKLAKAIGAQCYLECSALTQKGLKAVFDEAILTIFHPKKKKKRCSEGHSCCSII
3229	A	25	722	AISAGRSAKMQLKPMEINPEMLNKVLSRLGVAGQWRFVDVLGLEEESLGSVPAPACALLLLFPLTAQHENFRKKQIEELKGQEVSPKVYFMKQTIGNSCGTIGLIHAVANNQDKLGFEDGSVLKQFLSETEKMSPEDRAKCFEKNEAIQAAHDAVAQEGQCRVDDKVNFHFILFNNVDGHLYELDGRMPFPVNHGASSEDTLLKDAAKVCREFTEREQGEVRFSAVALCKAA
3230	A	282	1479	GDAATTACAPPDWFLGPRKLAAGPAGGGMLPRRLLAAWLAGTRGGGLLALLANQCRFVTGLRVRRAQQIAQLYGRLYSESSRRVLLGRLWRRLHGRPGHASALMAALAGVFVWDEERIQEEELQRSINEMKRLEEMSNMFQSSGVQHHPPEPKAQTEGNEDSEGKEQRWEMVMDKKHFKLWRRPITGTHLYQYRVFGTYTDVTPRQFFNVQLDTEYRKKWDALVIKLEVIERDVVSGSEVLHWVTHFPYPMYSRDYVYVRRYSVDQENNMMVLVSRAVEHPSVPESPEFVRVRSYESQMVIRPHKSFDENGFDYLLTYSDNPQTVFPR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YCVSWMVSSGMPDFLEKLHMATLKAKNMEIKV KDYISAKPLEMSSEAKATSQSSERKNEGSCGPAR IEYA
3231	A	2117	590	FVPEPPEAGASSPCAPGDPDMSFRKVVRQSKFRH VFGQPVKNDQCYEDIRVSRVTWDSTFCAVNPKF LAVIVEASGGGAFLVLPLSKTGRIDKAYPTVCGH TGPVLDIDWCPHNDEVIASGSEDCTVMVWQIPE NGLTSPLTEPVVVLEGHTKRVGIIAWHPTARNVL LSAGCDNVVLIWNVGTAEELYRLDSLHPDLIYN VSWNHNGSLFCSACKDKSVRIIDPRRGTLVAERE KAHEGARPMRAIFLADGKVFTTGFSRMSERQLA LWDPENLEEPMALQELDSSNGALLPFYDPDTSV VYVCGKGDSSIRYFEITEEPPYIHFLNTFTSKEPQR GMGSMPKRGLEVSKCEIARFYKLHERKCEPIVM TVPRKSDLFQDDLYPDTAGPEAALEAEEWVSGR DADPILISLREAYVPSKQRDLKISRRNVLSDSRPA MAPGSSHLGAPASTTTAADATPSGSLARAGEAG KLEEVMQELRALRALVKEQGDRICRLEEQLGRM ENGDA
3232	A	3	718	RLREDDRRGLPLSSPLWTEPPLSCCLPATYPADM GTAGAMQLCWVILGFLLFRGHNSQPTMTQTSSS QGGLGGLSLTTEPVSSNPGYIPSSEANRPSHLSST GTPGAGVPSSGRDGGTSRDTFQTVPPNSTTMSLS MREDATILPSPTSETVLTVAAFGVISFIVILVVVVI ILVGVVSLRFKCRKSKESEDPQKPGSSGLSESCST ANGEKDSITLISMKNINMNNGKQSLSAEKVL
3233	A	3	718	RLREDDRRGLPLSSPLWTEPPLSCCLPATYPADM GTAGAMQLCWVILGFLLFRGHNSQPTMTQTSSS QGGLGGLSLTTEPVSSNPGYIPSSEANRPSHLSST GTPGAGVPSSGRDGGTSRDTFQTVPPNSTTMSLS MREDATILPSPTSETVLTVAAFGVISFIVILVVVVI ILVGVVSLRFKCRKSKESEDPQKPGSSGLSESCST ANGEKDSITLISMKNINMNNGKQSLSAEKVL
3234	A	1169	4292	AGDCGRLGVGGSEFPWEGSALGASPLPPICLQSR TWLLRAPAPAELGELEEVAAGRGDVWEPFLDSP GREESLQEASPRLADHGSSSGGGWEVKRSQRLR RGPSSPRRPYQDMEYERRGGRGDRTGRYGATDR SQDDGGENRSRDHDYRDMDYRSYPREYGSQEG KHDYDDSSEEQSAEDSYEASPGSETQRRRRRRH RHSPTGPPGFPRDGDYRDQDYRTEQGEEEEEEED EEEEEKASNIVMLRMLPQAATEDDIRGQLQSHG VQAREVRLMRNKSSGQSRGFAFVEFSHLQDATR WMEANQHSLNILGQKVSMHYSDPKPKINEDWL CNKCGVQNFKRREKCFKCGVPKSEAEQKLPLGT RLDQQTLPLGGRELSQGLLPLPQPYQAQGVLAS QALSQGSEPSSENANDTIILRNLNPHSTMDSILGA LAPYAVLSSSNVRVIKDKQTQLNRGFAFIQLSTIE AAQLLQILQALHPPLTIDGKTINVEFAKGSKRDM ASNEGSRISAASVASTAIAAAQWAISQASQGGEG TWATSEEPPVDYSYYQQDEGYGNSQGTESSLYA HGYLKGTKGPGITGTKGDPTGAGPEASLEPGADS VSMQAFSRPQPGAAPGIYQQSAEASSSQGTAANS QSYTIMSPAVLKSELQSPTHPSSALPPATSPTAQE SYSQYPVPDVSTYQYDETSGYYYDPQTGLYYDP NSQYYYNAQSQQYLYWDGERRTYVPALEQSAD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GHKETGAPSKEGKEKKEKHKTKTAQQIAKDMERWARSLNKQKENFKNSFQPISSLRDDERRESATADAGYAILEKKGALAERQHTSMDLPKLASDDRPSPPRGLVAAYSGESDSEEEQERGGPEREEKLTDWQKLACLLCRRQFPSKEALIRHQQLSGLHKQNLEIHRRAHLSENELEALEKNDMEQMKYRDRAAERREKYGIPEPPEPKRRKYGGISTASVDFEQPTRDGLGSDNIGSRMLQAMGWKEGSGLGRKKQGIVTPIEAQTRVRGSGLGARGSSYGVTSTESYKETLHKTMVTRFNEAQ
3235	A	3	1217	PSFLNTGLGPTALGVLGGAGAGLMSNPSPQVPEEEASTSVCRPKSSMASTSRRQRRERRFRRYLSAGRLVRAQALLQRHPGLDVDAGQPPPLHRACARHDAPALCLLLRLGADPAHQDRHGDTALHAAARQGPDAYTDFFLPLLSRCPSAMGIKNKDGETPGQILGWGPPWDSAEEEEEDDASKEREWRQKLQGELEDEWQEVMGRFEGDASHETQEPESFSAWSDRLAREHAQKCQQQQREAEGSCRPPRAEGSSQSWRQQEEEQRLFRERARAKEEELRESRARRAQEALGDREPKPTRAGPREEHPRGAGRGSLWRFGDVPWPCPGGGDPEAMAAALVARGPPLEEQGALRRYLRVQQVRWHPDRFLQRFRSQIETWELGRVMGAVTALSQALNRHAEALK
3236	A	3	1416	GPASGMAEPTSDFETPIGWHASPELTPTLGPLSDTAPPRDRWMFWAMLPPPPPPLTSSLPAAGSKPSSESQPPMEAQSLPGAPPPFDAQILPGAQPPFDAQSPLDSQPQPSGQPWNFHASTSWYWRQSSDRFPRHQKSLNPAVKNSYYPRKYDAKFTDFSLPPSRKQKKKKRKEPVFHFFCDTCDRGFKNQEKYDKHMSEHTKCPELDCSFTAHEKIVQFHWRNMHAPGMKKIKLDTPEEIARWREERRKNYPTLANIERKKKLKLEKEKRGAVLTTTQYGKMKGMSRHSQMAKIRSPGKNHKWKNDNSRQRAVTGSGSHLCDLKLEGPPEANADPLGVLINSDSESDKEEKPQHSVIPKEVTPALCSLMSSYGSLSGSESEPEETPIKTEADVLAENQVLDSSAPKSPSQDVKATVRNFSEAKSENRKKSFEKTNPKREKRLSQLSNVIRTKNTPSISLGNASSSGHST
3237	A	3806	2204	FVGEQEGGCEAGAGRGAQTYPGEAGERWFGRRRRRGRVVSRKKMSLKSERRGIHVDQSDLLCKKGCGYYGNPAWQGFCSKCWREEYHKARQKQIQEDWELAERLQREEEEAFASSQSSQGAQSLTFSKFEEKKTNEKTRKVTTVKKFFSASSRVGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIEFLKTFHKTGQEIYKQTKLFLEGMHYKRDLSIEEQSECAQDFYHNVAERMQTRGKVPPERVEKIMDQIEKYIMTRLYKYVFCPETTDDEKKDLAIQKRIRALRWVTPQMLCVPVNEDIPEVSDMVVKAITDIIEMDSKRVPRDKLACITKCSKHIFNAIKITKNEPASADDFLPTLIYIVLKGNPPRLQSNIQYITRFCNPSRLMTGEDGYYFTNLCCAVAFIEKLDAQSLNLSQEDFDRYMSGQTSPRKQEAESWSPDACLGVKQMYKNLDLLSQLNERQERIMNEAKKLEKDLIDWTDGIAREVQDIVEKYPLEIKPPNQPLAAIDSENVENDKLPPPLQPQVYAG
3238	A	1373	449	VLSVCPTGVFRPAPCRMAFMKKYLLPILGLFMAYYYYSANEEFRPEMLQGKKVIVTGASKGIGREM

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AYHLAKMGAHVVVTARSKETLQKVVSHCLELG AASAHYIAGTMEDMTFAEQFVAQAGKLMGGLD MLILNHITNTSLNLFHDDIHHVRKSMEVNFLSYV VLTVAALPMLKQSNGSIVVVSSLAGKVAYPMVA AYSASKFALDGFFSSIRKEYSVSRVNVSITLCVLG LIDTETAMKAVSGIVHMQAAPKEECALEIIKGGA LRQEEVYYDSSLWTTLLIRNPCRKILEFLYSTSYN MDRFINK
3239	A	213	422	ERTMQLEIKVALNFIIFYLYNKLLW/QPLKKK*EA HWYPDKPLKGSGFHT/GEMVDPVGELAAKRSGL TVED
3240	A	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWLGA VAHSCNPSTLVGRGGRITRGQELR
3241	A	161	547	PAGIGRSTAKTPGTPGSLEMENLKSGVYPLKEAS GCPGADRNLLVYSFYEKGPLTFRDVAIEFSLEEW QCLDTAQQDLYRKVMLENYRNLVFLAGIAVSKP DLITCLEQGKEPWNMKRHAMVDQPPGR
3242	A	50	241	PLPARGKSTLPATFCSPSAPELASMSVVPPNRSQT GWPRGVTQFGNKYIQQTKPLTLERTINL
3243	A	380	702	FVAYLKLPFFSQVCLFASSEMFFTISRKNMSQKLS LLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQI LDLSKRYVKALAEENKNTVDVENGASMAGYGK ITVEYF
3244	A	37	1391	VLMDGRMMRSMRLREEESPGPSHTASCLCGSAP CILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSP GVESQLYKLPWVCEEGAGIPTVLQGHIDCGSLLG YRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRA AIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFY FGVVGSFLFILIQLVLLIDFAHSWNQRWLGKAEE CDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYT EPSGCHEGKVFISLNLTFCVCVSIAAVLPKVQDA QPNSGLLQASVITLYTMFVTWSALSSIPEQKCNP HLPTQLGNETVVAGPEGYETQWWDAPSIVGLIIF LLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQ QQQQVAACEGRAFDNEQDGVTYSYSFFHFCLVL ASLHVMMTLTNWYKPGETRKMISTWTAVWVKI CASWAGLLLYL
3245	A	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAHQV LTFLLLFVITSVASENASTSRGCGLDLLPQYVSLC DLDAIWGIVVEAAAGAGALITLLLMLILLVRLPF FKEKEKKSPVGLHFLFLLGTLGP
3246	A	3	515	HEVCGSGCCCHCCAGGPVARQKALPRLRGVMS RFLNVLRSWLVMVSIIAMGNTLQSFRDHTFLYEK LYTGKPNLVNGLQARTFGIWTLLSSVIRCLCAIDI HNKTLYHITLWTFLLALGHFLSELFVYGTAAPTI GVLAPLMVASFSILGMLVGLRYLEVEPVSRQKK RN
3247	A	1	932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEENSV THHEVKCQGKPLAGIYRKREEKRNAGNAVRSA MKSEEQKIKDARKGPLVPFPNQKSEAAEPPKTPP SSCDSTNAAIAKQALKKPIKGKQAPRKKAQGKT QQNRKLTDFYPVRRSSRKSKAELQSEERKRIDELI ESGKEEGMKIDLIDGKGRGVIATKQFSRGDFVVE YHGDLIEITDAKKREALYAQDPSTGCYMYYFQY LSKTYCVDATRETNRLGRLINHSKCGNCQTKLH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DIDGVPHLILIASRDIAAGEELLYDYGDRSKASIE AHPWLKH
3248	A	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWLRG VTATFGRPAEWPGYLSHLCGRSAAMDLGPMRK SYRGDREAFEETHLTSLDPVKQFAAWFEEAVQC PDIGEANAMCLATCTRDGKPSARMLLLKGFGKD GFRFFTNFESRKGKELDSNPFASLVFYWEPLNRQ VRVEGPVKKLPEEEAECYFHSRPKSSQIGAVVSH QSSVIPDREYLRKKNEELEQLYQDQEVPKPKSW GGYVLYPQVMEFWQGQTNRLHDRIVFRRGLPTG DSPLGPMTHRGEEDWLYERLAP
3249	A	43	1210	TRVGRGESGLKMEVKPPPGRPQPDSGRRRRRRG EEGHDPKEPEQLRKLFIGGLSFETTDDSLREHFEK WGTLTDCVVMRDPQTKRSRGFGFVTYSCVEEV DAAMCARPHKVDGRVVEPKRAVSREDSVKPGA HLTVKKIFVGGIKEDTEEYNLRDYFEKYGKIETIE VMEDRQSGKKRGFAFVTFDDHDTVDKIVVQKY HTINGHNCEVKKALSKQEMQSAGSQRGRGGGS GNFMGRGGNFGGGGGNFGRGGNFGGRGGYGG GGGGSRGSYGGGDGGYNGFGGDGGNYGGGPG YSSRGGYGGGGPGYGNQGGGYGGGGGYDGYN EGGNFGGGNYGGGGNYNDFGNYSGQQQSNYGP MKGGSFGGRSSGSPYGGGYGSGGGSGGYGSRRF
3250	A	32	1175	VAGRGDMAALRDAEIQKDVQTYYGQVLKRSAD LQTNGCVTTARPVPKHIREALQNVHEEVALRYY GCGLVIPEHLENCWILDLGSGSGRDCYVLSQLVG EKGHVTGIDMTKGQVEVAEKYLDYHMEKYGFQ ASNVTFIHGYIEKLGEAGIKNESHDIVVSNCVINL VPDKQQVLQEAYRVLKHGGELYFSDVYTSLELP EEIRTHKVLWGECLGGALYWKELAVLAQKIGFC PPRLVTANLITIQNKELERVIGDCRFVSATFRLFK HSKTGPTKRCQVIYNGGITGHEKELMFDANFTFK EGEIVEVDEETAAILKNSRFAQDFLIRPIGEKLPTS GGCSALELKDIITDPFKLAEESDSMKSRCVPDAA GGCCGTKKSC
3251	A	32	1175	VAGRGDMAALRDAEIQKDVQTYYGQVLKRSAD LQTNGCVTTARPVPKHIREALQNVHEEVALRYY GCGLVIPEHLENCWILDLGSGSGRDCYVLSQLVG EKGHVTGIDMTKGQVEVAEKYLDYHMEKYGFQ ASNVTFIHGYIEKLGEAGIKNESHDIVVSNCVINL VPDKQQVLQEAYRVLKHGGELYFSDVYTSLELP EEIRTHKVLWGECLGGALYWKELAVLAQKIGFC PPRLVTANLITIQNKELERVIGDCRFVSATFRLFK HSKTGPTKRCQVIYNGGITGHEKELMFDANFTFK EGEIVEVDEETAAILKNSRFAQDFLIRPIGEKLPTS GGCSALELKDIITDPFKLAEESDSMKSRCVPDAA GGCCGTKKSC
3252	A	1	574	PLGSNTAPALRVMVQAWYMDDAPGDPRQPHRP DPGRPVGLEQLRRLGVLYWKLDADKYENDPELE KIRRERNYSWMDIITICKDKLPNYEEKIKMFYEE HLHLDDEIRYILDGSGYFDVRDKEDQWIRIFMEK GDMVTLPAGIYHRFTVDEKNYTKAMRLFVGEPV WTAYNRPADHFEARGQYVKFLAQTA
3253	A	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPVLLA SLGVGLVTLLGLAVGSYLVRRSRRPQVTLLDPNE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KYLLRLLDKTTVSHNTKRFRFALPTAHHTLGLPVGKHIYLSTRIDGSLVIRPYTPVTSDEDQGYVDLVIKVYLKGVHPKFPEGGKMSQYLDSLKVGDVVEFRGPSGLLTYTGKGHFNIQPNKKSPPEPRVAKKLGMIAGGTGITPMLQLIRAILKVPEDPTQCFLLFANQTEKDIILREDLEELQARYPNRFKLWFTLDHPPKDWAYSKGFVTADMIREHLPAPGDDVLVLLCGPPPMVQLACHPNLDKLGYSQKMRFTY
3254	A	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRRRYHRLSDMSMLAERRRKQKWAVDPQNTAWSNDDSKFGQRMLEKMGWSKGKGLGAQEQGATDHIKVQVKNNHLGLGATINNEDNWIAHQDDFNQLLAELNTCHGQETTDSSDKKEKKSFSLEEKSKISKNRVHYMKFTKGKDLSSRSKTDLDCIFGKRQSKKTPEGDASPSTPEENETTTTSAFTIQEYFAKRMAALKNKPQVPVPGSDISETQVERKRGKKRNKEATGKDVESYLQPKAKRHTEGKPERAEAQERVAKKKSAPAEEQLRGPCWDQSSKASAQDAGDHVQPA
3255	A	173	439	GSAAMKVKIKCWNGVATWLWVANDENCGICRMAFNGCCPDCKVPGDDCPLVWGQCSHCFHMHCILKWLHAQQVQQHCPMCRQEWKFKE
3256	A	2	377	TAARRRQKGTAARRRQKGTLEEVVLPPRSCRVFWIHSGTTMSKVSFKITLTSDPRLPYKVLSVPESTPFTAVLKFAAEEFKVPAATSAIITNDGIGINPAQTAGNVFLKHGSELRIIPRDRVGSC
3257	A	3	1454	GCSAAAAGAGSGPWAAQEKQFPPALLSFFIYNPRFGPREGQEENKILFYHPNEVEKNEKIRNVGLCEAIVQFTRTFSPSKPAKSLHTQKNRQFFNEPEENFWMVMVVRNPIIEKQSKDGKPVIEYQEEELLDKVYSSVLRQCYSMYKLFNGTFLKAMEDGGVKLLKERLEKFFHRYLQTLHLQSCDLLDIFGGISFFPLDKMTYLKIQSFINRMEESLNIVKYTAFLYNDQLIWSGLEQDDMRILYKYLTTSLFPRHIEPELAGRDSPIRAEMPGNLQHYGRFLTGPLNLNDPDAKCRFPKIFVNTDDTYEELHLIVYKAMSAAVCFMIDASVHPTLDFCRRLDSIVGPQLTVLASDICEQFNINKRMSGSEKEPQFKFIYFNHMNLAEKSTVHMRKTPSVSLTSVHPDLMKILGDINSDFTRVDEDEIIVKAMSDYWVVGKKSDRRELYVILNQKNANLIEVNEEVKKLCATQFNNIFFLD
3258	A	113	1558	APRGCSMPHRKKKPFIEKKKAVSFHLVHRSQRDPLAADESAPQRVLLPTQKIDNEERRAEQRKYGVFFDDDYDYLQHLKEPSGPSELIPSSTFSAHNRREEKEETLVIPSTGIKLPSSVFASEFEEDVGLLNKAAPVSGPRLDFDPDIVAALDDDFDFDDPDNLLEDDFILQANKATGEEEGMDIQKSENEDDSEWEDVDDEKGDSNDDYDSAGLLSDEDCMSVPGKTHRAIADHLFWSEETKSRFTEYSMTSSVMRRNEQLTLHDERFEKFYEQYDDDEIGALDNAELEGSIQVDSNRLQEVLNDYYKEKAENCVKLNTLEPLEDQDLPMNELDESEEEEMITVVLEEAKEKWDCESICSTYSNLYNHPQLIKYQPKPKQIRISSKTGIPLNVLPKKGLTAKQTERIQMINGSDLPKVSTQPRSKNESKEDKRARKQAIKEERKERRVEKKANKLAFKLEKRRQEKELLNLKKNVEGLKL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3259	A	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFLSMYLVTVLGNLLIILATISDSHLHTPMYFFLSNLSFADICVTSTTIPKMLMNIQTQNKVITYIACLMQMYFFILFAGFENFLLSVMAYDRFVAICHPLHYMVIMNPHLCGLLVLASWTMSALYSLLQILMVVRLSFCTALEIPHFFCELNQVIQLACSDSFLNHMVIYFTVALLGGGPLTGILYSYSKIISSIHAISSAQGKYKAFSTCASHLSVVSLFYGAILGVYLSSAATRNSHSSATASVMYTVVTPMLNPFIYSLRNKDIKRALGIHLLWGTMKGQFFKKCP
3260	A	34	2573	IPFLKSCCCCCLFDFPPPPLDQVQEEECEVERVTEHGTPKPFRKFDSVAFGESQSEDEQFENDLETDPPNWQQLVSREVLLGLKPCEIKRQEVINELFYTERAHVRTLKVLDQVFYQRVSREGILSPSELRKIFSNLEDILQLHIGLNEQMKAVRKRNETSVIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQPFALEMIKSRQKKDSRFQTFVQDAESNPLCRRLQLKDIIPTQMQRLTKYPLLLDNIATYTEWPTEREKVKKAADHCRQILNYVNQAVKEAENKQRLEDYQRRLDTSSLKLSEYPNVEELRNLDLTKRKMIHEGPLVWKVNRDKTIDLYTLLLEDILVLLQKQDDRLVLRCHSKILASTADSKHTFSPVIKLSTVLVRQVATDNKALFVISMSDNGAQIYELVAQTVSEKTVWQDLICRMAASVKEQSTKPIPLPQSTPGEGDNDEEDPSKLKEEQHGISVTGLQSPDRDLGLESTLISSKPQSHSLSTSGKSEVRDLFVAERQFAKEQHTDGTLKEVGEDYQIAIPDSHLPVSEERWALDALRNLGLLKQLLVQQLGLTEKSVQEDWQHFPRYRTASQGPQTDSVIQNSENIKAYHSGEGHMPFRTGTGDIATCYSPRTSTESFAPRDSVGLAPQDSQASNILVMDHMIMTPEMPTMEPEGGLDDSGEHFFDAREAHSDENPSEGDGAVNKEEKDVNLRISGNYLILDGYDPVQESSTDEEVASSLTLQPMTGIPAVESTHQQQHSPQNTHSDGAISPFTPEFLVQQRWGAMEYSCFEIQSPSSCADSQSQIMEYIHKIEADLEHLKKVEESYTILCQRLAGSALTDKHSDKS
3261	A	1	2100	AVEFAEGALTMAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLIDEPTEVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLLGFLYFFVCSLDILSSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIPIIMGANIGTSITNTIVALMQVGDRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEIITQLIVESFHFKNGEDAPDLLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKTFTNKTQINVTVPSTANCTSPSLCWTDGIQNWTMKNVTYKENIAKCQHIFVNFHLPDLAVGTILLILSLLVLCGCLIMIVKILGSVLKGQVATVIKKTINTDFPFPFAWLTGYLAILVGAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTLGSNIGTTTTAILAALASPGNALRSSLQIALCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIPLTVFGLSLAGWRVLVGVGVPVVFIIILVLCLRLLQSRCPRVLPKKLQNWNFLPLWMRSLKPWDAVVSKFTGCFQMRCCCCCRVCCRACCLLCGCPKCCRCSKCCEDLEEAQEGQDVPVKAPETFDNITISREAQGEVPA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SDSKTECTAL
3262	A	30	1377	SQQGSQPHRQGPPSLLTAPHSLDLPALPPGPRGS QGKLRRVLVPMSVKPSWGPGPSEGVTAVPTSDL GEIHNWTELLDLFNHTLSECHVELSQSTKRVVLF ALYLAMFVVGLVENLLVICVNWRGSGRAGLMN LYILNMAIADLGIVLSLPVWMLEVTLDYTWLWG SFSCRFTHYFYFVNMYSSIFFLVCLSVDRYVTLTS ASPSWQRYQHRVRRAMCAGIWVLSAIIPLPEVV HIQLVEGPEPMCLFMAPFETYSTWALAVALSTTI LGFLLPFPLITVFNVLTACRLRQPGQPKSRRHCLL LCAYVAVFVMCWLPYHVTLLLLTLHGTHISLHC HLVHLLYFFYDVIDCFSMLHCVINPILYNFLSPHF RGRLLNAVVHYLPKDQTKAGTCASSSSCSTQHSI IITKGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP TQPLTPS
3263	A	1	919	QARSPSVAAMASPQLCRALVSAQWVAEALRAP RAGQPLQLLDASWYLPKLGRDARREFEERHIPG AAFFDIDQCSDRTSPYDHMLPGAEHFAEYAGRL GVGAATHVVIYDASDQGLYSAPRVWWMFRAFG HHAVSLLDGGLRHWLRQNLPLSSGKSQPAPAEF RAQLDPAFIKTYEDIKENLESRRFQVVDSRATGR FRGTEPEPRDGIEPGHIPGTVNIPFTDFLSQEGLEK SPEEIRHLFQEKKVDLSKPLVATCGSGVTACHVA LGAYLCGKPDVPIYDGSWVEWYMRARPEDVISE GRGKTH
3264	A	1	1398	ARRSTPRTAPRASATRSAAGTMREIVHIQAGQCG NQIGAKFWEVISDEHGIDPTGSYHGDSDLQLERI NVYYNEAAGNKYVPRAILVDLEPGTMDSVRSGP FGQIFRPDNFVFGQSGAGNNWAKGHYTEGAELV DSVLDVVRKESESCDCLQGFQLTHSLGGGTGSG MGTLLISKIREEYPDRIMNTFSVMPSPKVSDTVVE PYNATLSVHQLVENTDETYSIDNEALYDICFRTL KLTTPTYGDLNHLVSATMSGVTTCLRFPGQLNA DLRKLAVNMVPFPRLHFFMPGFAPLTSRGSQQY RALTVPELTQQMFDSKNMMAACDPRHGRYLTV AAIFRGRMSMKEVDEQMLNVQNKNSSYFVEWIP NNVKTAVCDIPPRGLKMSATFIGNSTAIQELFKRI SEQFTAMFRRKAFLHWYTGEGMDEMEFTEAES NMNDLVSEYQQYQDATADEQGEFEEEEGEDEA
3265	A	265	862	WWEDARVLGPFHPEEEGHWVMTPSEGARAGTG RELEMLDSLLALGGLVLLRDSVEWEGRSLLKAL VKKSALCGEQVHILGCEVSEEEFREGFDSDINNR LVYHDFFRDPLNWSKTEEAFPGGPLGALRAMCK RTDPVPVTIALDSLSWLLLRLPCTTLCQVLHAVS HQDSCPGETPPSLFPLIHLPLPRSVPLFLSTLE
3266	A	2	884	AAGAGADGREPASERASRAEPPAVAMGQNDLM GTAEDFADQFLRVTKQYLPHVARLCLISTFLEDG IRMWFQWSEQRDYIDTTWNCGYLLASSFVFLNL LGQLTGCVLVLSRNFVQYACFGLFGIIALQTIAYS ILWDLKFLMRNLALGGGLLLLLAESRSEGKSMF AGVPTMRESSPKQYMQLGGRVLLVLMFMTLLH FDASFFSIVQNIVGTALMILVAIGFKTKLAALTLV VWLFAINVYFNAFWTIPVYKPMHDFLKYDFFQT MSVIGGLLLVVALGPGGVSMDEKKKEW
3267	A	802	1011	ASTFCSAWKRRSTAALWWSGSRASRSHPRELGP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LCFVFGTAALSIRSMDVLSLFLEHGKLVFASGLSPRA
3268	A	490	679	EDAWITNPSLSNARSTPSKPLCYTVLKEGQVVGVKTTKASNTREKLRPESERRMVKSFGDEVT
3269	A	2	796	GSTHASGARPSLKRARSQRGRPLPSRALPSAHKDMTTNAGPLHPYWPQHLRLDNFVPNDRPTWHILAGLFSVTGVLVVTTWLLSGRAAVVPLGTWRRLSLCWFAVCGFIHLVIEGWFVLYYEDLLGDQAFLSQLWKEYAKGDSRYILGDNFTVCMETITACLWGPLSLWVVIAFLRQHPLRFILQLVVSVGQIYGDVLYFLTEHRDGFQHGELGHPLYFWFYFVFMNALWLVLPGVLVLDAVKHLTHAQSTLDAKATKAKSKKN
3270	A	17	229	GDTGPQILMSYLDSVASKLLQMVKKLSQSFCSNFKYLTKYSRKQVSDEIKKSRRTVESNPIFFKKNKKIQ
3271	A	419	553	IQSGLSLCFADLSETPEGRAGVPGCPHSCDGVASGRPCSPSSAG
3272	A	1211	1450	FQFIQIELLNILQSLIRNQTQSPYNTTAYPAIDSVITILPFSFSCFFIITKCFGLSIFPSVIFFLHVYFILTLVVFYCC
3273	A	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPQLLFPELPLPHVPGQESAKRRSARRFLIMSELTKELMELVWGTKSSPGLSDTIFCRWTQGFVFSESEGSALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSSWRDCSQEEQKELLCHTLCDILESACCDHSGSYCLVSWLRGKTTEETASISGSPAESSCQVEHSSALAVEELGFERFHALIQKRSFRSLPELKDAVLDQYSMWGNKFGVLLFLYSVLLTKGIENIKNEIEDASEPLIDPVYGHGSQSLINLLLTGHAVSNVWDGDRECSGMKLLGIHEQAAVGFLTLMEALRYCKVGSYLKISKIPYLDCLASETHLTVFFAKDMALVAPEAPSEQARRVFQTYDPEDNGFIPDSLLEDVMKALDLVSDPEYINLMKNKLDPEGLGIILLGPFLQEFFPDQGSSGPESFTVYHYNGLKQSNYNEKVMYVEGTAVVMGFEDPMLQTDDTPIKRCLQTKWPYIELLWTTDRSPSLN
3274	A	186	1358	RVVHRFFKSSAFWPAEVKQPRGGPKTGSRKEGAGSRAPQPVVRSFCGSVGAEGRMEKLRLLGLRYQEYVTRHPAATAQLETAVRGFSYLLAGRFADSHELSELVYSASNLLVLLNDGILRKELRKKLPVSLSQQKLLTWLSVLECVEVFMEMGAAKVWGEVGRWLVIALIQLAKAVLRMLLLLWFKAGLQTSPPIVPLDRETQAQPPDGDHSPGNHEQSYVGKRSNRVVRTLQNTPSLHSRHWGAPQQREGRQQQHHEELSATPTPLGLQETIAEFLYIARPLLHLLSLGLWGQRSWKPWLLAGVVDVTSLSLLSDRKGLTRRERRELRRRTILLLYYLLRSPFYDRFSEARILFLLQLLADHVPGVGLVTRPLMDYLPTWQKIYFYSWG
3275	A	575	759	SVYSASSCKCCNYRKTEQIPDCEQPPASSMPERPSHESQPTPQMMPLSAPSRAEELGQRPG
3276	A	7	258	KAAGHRLLLAAGHPSMPSSDCLLWEGSLELRPLQHISSLLVLVSTTCLFAFPRVPIAFESKSCLIYHCHCAFTVRHYMCSSHTG
3277	A	9	2221	KLGVEPEEEGGGDDEEDAEAWAMELADVGAAASSQGVHDQVLPTPNASSRVIVHVDLDCFYAQVEMISNPELKDKPLGVQQKYLVVTCNYEARKLGVK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KLMNVRDAKEKCPQLVLVNGEDLTRYREMSYK VTELLEEFSPVVERLGFDENFVDLTEMVEKRLQQ LQSDELSAVTVSGHVYNNQSINLLDVLHIRLLVG SQIAAEMREAMYNQLGLTGCAGVASNKLLAKL VSGVFKPNQQTVLLPESCQHLIHSLNHIKEIPGIG YKTAKCLEALGINSVRDLQTFSPKILEKELGISVA QRIQKLSFGEDNSPVILSGPPQSFSEEDSFKKCSSE VEAKNKIEELLASLLNRLCQDERKPHTVRLIIRRY SSEKHYGRESRQCPIPSHVIQKLGTGNYDVMTPM VDILMKLFRNMVNVKMPFHLTLLSVCFCNLKAL NTAKKGLIDYYLMPSLSTTSRSGKHSFKMKDTH MEDFPKDKETNRDFLPSGRIESTRTRESPLDTTNF SKEKDINEFPLCSLPEGVDQEVFKQLPVDIQEEIL SGKSREKFQGKGSVSCPLHASRGVLSFFSKKQM QDIPINPRDHLSSSKQVSSVSPCEPGTSGFNSSSSS YMSSQKDYSYYLDNRLKDERISQGPKEPQGFHF TNSNPAVSAFHSFPNLQSEQLFSRNHTTDSHKQT VATDSHEGLTENREPDSVDEKITFPSDIDPQVFYE LPEAVQKELLAEWKRTGSDFHIGHK
3278	A	1	876	GLRLHVDLVEKPRTGIMAAETRNVAGAEAPPPQ KRYYRQRAHSNPMADHTLRYPVKPEEMDWSEL YPEFFAPLTQNQSHDDPKDKKEKRAQAQVEFAD IGCGYGGLLVELSPLFPDTLILGLEIRVKVSDYVQ DRIRALRAAPAGGFQNIACLRSNAMKHLPNFFY KGQLTKMFFLFPDPHFKRTKHKWRIISPTLLAEY AYVLRVGGLVYTITDVLELHDWMCTHFEEHPLF ERVPLEDLSEDPVVGHLGTSTEEGKKVLRNGGK NFPAIFRRIQDPVLQAVTSQTSLPGH
3279	A	82	2929	TRTKRRLGREKAMASPPRGWGCGELLLPFMLLG TLCEPGSGQIRYSMPEELDKGSFVGNIAKDLGLE PQELAERGVRIVSRGRTQLFALNPRSGSLVTAGRI DREELCAQSPLCVVNFNILVENKMKIYGVEVEII DINDNFPRFRDEELKVKVNENAAAGTRLVLPFA RDADVGVNSLRSYQLSSNLHFSLDVVSGTDGQK YPELVLEQPLDREKETVHDLLLTALDGGDPVLSG TTHIRVTVLDANDNAPLFTPSEYSVSVPENIPVGT RLLMLTATDPDEGINGKLTYSFRNEEEKISETFQL DSNLGEISTLQSLDYEESRFYLMEVVAQDGGAL VASAKVVVTVQDVNDNAPEVILTSLTSSISEDCL PGTVIALFSVHDGDSGENGEIACSIPRNLPFKLEK SVDNYYHLLTTRDLDREETSDYNITLTVMDHGT PPLSTESHIPLKVADVNDNPPNFPQASYSTSVTEN NPRGVSIFSVTAHDPDSGDNARVTYSLAEDTFQG APLSSYVSINSDTGVLYALRSFDYEQLRDLQLWV TASDSGNPPLSSNVSLSLFVLDQNDNTPEILYPAL PTDGSTGVELAPRSAEPGYLVTKVVAVDKDSGQ NAWLSYRLLKASEPGLFAVGLHTGEVRTARALL DRDALKQSLVVAVEDHGQPPLSATFTVTVAVAD RIPDILADLGSIKTPIDPEDLDLTLYLVVAVAAVS CVFLAFVIVLLVLRLRRWHKSRLLQAEGSRLAG VPASHFVGVDGVRAFLQTYSHEVSLTADSRKSH LIFPQPNYADTLLSEESCEKSEPLLMSDKVDANK EERRVQQAPPNTDWRFSQAQRPGTSGSQNGDDT GTWPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYRQN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VYIPGSNATLTNAAGKRDGKAPAGGNGNKKKSGKKEKK
3280	A	149	1288	GTSQMSSHKGSVVAQGNGAPASNREADTAELAELGPLLEEKGKRVIANPPKAEEEQTCPVPQEEEEEVRVLTLPLQAHHAMEKMEEFVYKVWEGRWRVIPYDVLPDWLKDNDYLLHGHRPPMPSFRACFKSIFRIHTETGNIWTHLLGFVLFLFLGILTMLRPNMYFMAPLQEKVVFGMFFLGAVLCLSFSWLFHTVYCHSEKVSRTFSKLDYSGIALLIMGSFVPWLYYSFYCSPQPRLIYLSIVCVLGISAIIVAQWDRFATPKHRQTRAGVFLGLGLSGVVPTMHFTIAEGFVKATTVGQMGWFFLMAVMYITGAGLYAARIPERFFPGKFDIWFQSHQIFHVLVVAAAFVHFYGVSNLQEFRYGLEGGCTDDTLL
3281	A	1	557	RPRRRQPSFSCRVLVLEDPPCFRFTNSMNQEKLAKLQAQVRIGGKGTARRKKKVVHRTATADDKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVIHFNNPKVQASLSANTFAITGHAEAKPITEMLPGILSQLGADSLTSLRKLAEQFPRQVLDSKAPKPEDIDEEDDDVPDLVENFDEASKNEAN
3282	A	155	1139	HALGRRGGSQELSAAACGCFALRLRAPGSGRPALAPGAAAFAGLGGAPRFPPRGSAAGRTMLLKEYRICMPLTVDEYKIGQLYMISKHSHEQSDRGEGVEVVQNEPFEDPHHGNGQFTEKRVYLNSKLPSWARAVVPKIFYVTEKAWNYYPYTITEYTCSFLPKFSIHIETKYEDNKGSNDTIFDNEAKDVEREVCFIDIACDEIPERYYKESEDPKHFKSEKTGRGQLREGWRDSHQPIMCSYKLVTVKFEVWGLQTRVEQFVHKVVRDILLIGHRQAFAWVDEWYDMTMDDVREYEKNMHEQTNIKVCNQHSSPVDDIESHAQTST
3283	A	159	547	IKSKLNQQVEVQESEWRLTEAKGPTMGKESGWDSGRAAVAAVVGGVVAVGTVLVALSAMGFTSVGIAASSIAAKMMSTAAIANGGGVAAGSLVAILQSVGAAGLSVTSKVIGGFAGTALGAWLGSPPSS
3284	A	227	637	TSNSLLRPDRMSVMDLANTCSSFQSDLDFCSDCGSVLPLPGAQDTVTCIRCGFNINVRDFEGKVVKTSVVFHQLGTAMPMSVEEGPECQGPVVDRRCPRCGHEGMAYHTRQMRSADEGQTVFYTCTNCKFQEKEDS
3285	A	123	1535	HRLSYDEAFAMANDPLEGFHEVNLASPTSPDLLGVYESGTQEQTTSPSVIYRPHPSALSSVPIQANALDVSELPTQPVYSSPRRLNCAEISSISFHVTDPAPCSTSGVTAGLTKLTTRKDNYNAEREFLQGATITEACDGSDDIFGLSTDSLSRLRSPSVLEVREKGYERLKEELAKAQRELKLKDEECERLSKVRDQLGQELEELTASLFEEAHKMVREANIKQATAEKQLKEAQGKIDVLQAEVAALKTLVLSSSPTSPTQEPLPGGKTPFKKGHTRNKSTSSAMSGSHQDLSVIQPIVKDCKEADLSLYNEFRLWKDEPTMDRTCPFLDKIYQEDIFPCLTFSKSELASAVLEAVENNTLSIEPVGLQPIRFVKASAVECGGPKKCALTGQSKSCKHRIKLGDSSNYYYISPFCRYRITSVCNFFTYIRYIQQGLVKQQDVDQMFWEVMQLRKEMSLAKLGYFKEEL
3286	A	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHGYPGITEELLRSQLYPEVPPEEFRPFLAKMRGILKSIAS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ADMDFNQLEAFLTAQTKKQGGITSDQAAVISKF WKSHKTKIRESLMNQSRWNSGLRGLSWRVDGK SQSRHSAQIHTPVAIIELELGKYGQESEFLCLEFD EVKVNQILKTLSEVEESISTLISQPN
3287	A	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDGKC VICDSYVRPCTLVRICDECNYGSYQGRCVICGGP GVSDAYYCKECTIQEKDRDGCPKIVNLGSSKTDL FYERKKYGFKKR
3288	A	3	428	RTTFFRFRPCESLCGDMKLLTHNLLSSHVRGVGS RGFPLRLQATEVRICPVEFNPNFVARMIPKVEWS AFLEAADNLRLIQVPKGPVEGYEENEEFLRTMH HLLLEVEVIEGTLQCPESGRMFPISRGIPNMLLSE EETES
3289	A	1	1743	AGCCRDTRFPTPRGPGSLCHNFCRSAACTVTRTI HGSPREDTGTPRSREMMFQDSVAFEDVAVSFTQ EEWALLDPSQKNLYRDVMQETFKNLTSVGKTW KVQNIEDEYKNPRRNLSLMREKLCESKESHHCG ESFNQIADDMLNRKTLPGITPCESSVCGEVGTGH SSLNTHIRADTGHKSSEYQEYGENPYRNKECKK AFSYLDSFQSHDKACTKEKPYDGKECTETFISHS CIQRHRVMHSGDGPYKCKFCGKAFYFLNLCLIH ERIHTGVKPYKCKQCGKAFTRSTTLPVHERTHTG VNADECKECGNAFSFPSEIRRHKRSHTGEKPYEC KQCGKVFISFSSIQYHKMTHTGEKPYECKQCGK AFRCGSHLQKHGRTHTGEKPYECRQCGKAFRCT SDLQRHEKTHTEDKPYGCKQCGKGFRCASQLQI HERTHSGEKPHECKECGKVFKYFSSLRIHERTHT GEKPHECKQCGKAFRYFSSLHIHERTHTGDKPYE CKVCGKAFTCSSSIRYHERTHTGEKPYECKHCGK AFISNYIRYHERTHTGEKPYQCKQCGKAFIRASS CREHERTHTINR
3290	A	2	1350	GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSAAS RRSPAARPPVPAPPALPRGRPGTEGSTSLSAPAVL VVAVAVVVVVVVSAVAWAMANYIHVPPGSPEVP KLNVTVQDQEEHRCREGALSLLQHLRPHWDPQE VTLQLFTDGITNKLIGCYVGNTMEDVVLVRIYGN KTELLVDRDEEVKSFRVLQAHGCAPQLYCTFNN GLCYEFIQGEALDPKHVCNPAIFRLIARQLAKIHA IHAHNGWIPKSNLWLKMGKYFSLIPTGFADEDIN KRFLSDIPSSQILQEEMTWMKEILSNLGSPVVLCH NDLLCKNIIYNEKQGDVQFIDYEYSGYNYLAYDI GNHFNEFAGVSDVDYSLYPDRELQSQWLRAYLE AYKEFKGFGTEVTEKEVEILFIQVNQFALASHFF WGLWALIQAKYSTIEFDFLGYAIVRFNQYFKMK PEVTALKVPE
3291	A	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASAQD ARYGQKDSSDQNFDYMFKLLIIGNSSVGKTSFLF RYADDSFTSAFVSTVGIDFKVKTVFKNEKRIKLQI WDTAGQERYRTITTAYYRGAMGFILMYDITNEE SFNAVQDWSTQIKTYSWDNAQVILVGNKCDME DERVISTERGQHLGEQLGFEFFETSAKDNINVKQ TFERLVDIICDKMSESLETDPAITAAKQNTRLKET PPPPQPNCAC
3292	A	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTSLQ QRTPAEMSPVLHFYVRPSGHEGAASGHTRRKLQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GKLPELQGVETELCYNVNWTAEALPSAEETKKL MWLFGCPLLLDDVARESWLLPGSNDLLLEVGPR LNFSTPTSTNIVSVCRATGLGPVDRVETTRRYRLS FAHPPSAEVEAIALATLHDRMTEQHFPHPIQSFSP ESMPEPLNGPINILGEGRLALEKANQELGLALDS WDLDFYTKRFQELQRNPSTVEAFDLAQSNSEHS RHWFFKGQLHVDGQKLVHSLFESIMSTQESSNP NNVLKFCDNSSAIQGKEVRFLRPEDPTRPSRFQQ QQGLRHVVFTAETHNFPTGVCPFSGATTGTGGRI RDVQCTGRGAHVVAGTAGYCFGNLHIPGYNLP WEDLSFQYPGNFARPLEVAIEASNGASDYGNKF GEPVLAGFARSLGLQLPDGQRREWIKPIMFSGGI GSMEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEQKM NRVIRACVEAPKGNPICSLHDQGAGGNGNVLKE LSDPAGAIIYTSRFQLGDPTLNALEIWGAEYQESN ALLLRSPNRDFLTHVSARERCPACFVGTITGDRRI VLVDDRECPVRRNGQGDAPPTPPPTPVDLELEW VLGKMPRKEFFLQRKPPMLQPLALPPGLSVHQA LERVLRLPAVASKRYLTNKVDRSVGGLVAQQQC VGPLQTPLADVAVVALSHEELIGAATALGEQPV KSLLDPKVAARLAVAEALTNLVFALVTDLRDVK CSGNWMWAAKLPGEGAALADACEAMVAVMA ALGVAVDGGKDSLSMAARVGTETVRAPGSLVIS AYAVCPDITATVTPDLKHPEGRGHLLYVALSPG QHRLGGTALAQCFSQLGEHPPDLDLPENLVRAFS ITQGLLKDRLLCSGHDVSDGGLVTCLLEMAFAG NCGLQVDVPVPRVDVLSVLFAEEPGLVLEVQEP DLAQVLKRYRDAGLHCLELGHTGEAGPHAMVR VSVNGAVVLEEPVGELRALWEETSFQLDRLQAE PRCVAEEERGLRERMGPSYCLPPTFPKASVPREP GGPSPRVAILREEGSNGDREMADAFHLAGFEVW DVTMQDLCSGAIGLDTFRGVAFVGGFSYADVLG SAKGWAAAVTFHPRAGAELRRFRKRPDTFSLGV CNGCQLLALLGWVGGDPNEDAAEMGPDSQPAR PGLLLRHNLSGRYESRWASVRVGPGPALMLRG MEGAVLPVWSAHGEGYVAFSSPELQAQIEARGL APLHWADDDGNPTEQYPLNPNGSPGGVAGICSC DGRHLAVMPHPERAVRPWQWAWRPPPFDTLTT SPWLQLFINARNWTLEGSC
3293	A	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRERV AMHYQMSVTLKYEIKKLIYVHLVIWLLLVAKMS VGHLRLLSHDQVAMPYQWEYPYLLSILPSLLGLL SFPRNNISYLVLSMISMGLFSIAPLIYGSMEMFPA AQQLYRHGKAYRFLFGFSAVSIMYLVLVLAVQV HAWQLYYSKKLLDSWFTSTQEKKHK
3294	A	35	1821	SQRSCPRSPSSPAPPWARCSNPDSRTGGVPVPRA WSAGGPALGLMAAPVRLGRKRPLPACPNPLFVR WLTEWRDEATRSRHRTRFVFQKALRSLRRYPLP LRSGKEAKILQHFGDGLCRMLDERLQRHRTSGG DHAPDSPSGENSPAPQGRLAEVQDSSMPVPAQP KAGGSGSYWPARHSGARVILLVLYREHLNPNGH HFLTKEELLQRCAQKSPRVAPGSARPWPALRSLL HRNLVLRTHQPARYSLTPEGLELAQKLAESEGLS LLNVGIGPKEPPGEETAVPGAASAELASEAGVQQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QPLELRPGEYRVLLCVDIGETRGGGHRPELLREL QRLHVTHTVRKLHVGDFVWVAQETNPRDPANP GELVLDHIVERKRLDDLCSSIIDGRFREQKFRLKR CGLERRVYLVEEHGSVHNLSLPESTLLQAVTNTQ VIDGFFVKRTADIKESAAYLALLTRGLQRLYQGH TLRSRPWGTPGNPESGAMTSPNPLCSLLTFSDFN AGAIKNKAQSVREVFARQLMQVRGVSGEKAAA LVDRYSTPASLLAAYDACATPKEQETLLSTIKCG RLQRNLGPALSRTLSQLYCSYGPLT
3295	A	2	1115	EFHPHTQVSGLLTPQLQEPDVWSPSRGQPVSLHL PGKGAPEVKEMAWWKSWIEQEGVTVKSSSHFN PDPDAETLYKAMKGIGTNEQAIIDVLTKRSNTQR QQIAKSFKAQFGKDLTETLKSELSGKFERLIVAL MYPPYRYEAKELHDAMKGLGTKEGVIIEILASRT KNQLREIMKAYEEDYGSSLEEDIQADTSGYLERI LVCLLQGSRDDVSSFVDPALALQDAQDLYAAGE KIRGTDEMKFITILCTRSATHLLRVFEEYEKIANK SIEDSIKSETHGSLEEAMLTVVKCTQNLHSYFAE RLYYAMKGAGTRDGTLIRNIVSRSEIDLNLIKCH FKKMYGKTLSSMIMEDTSGDYKNALLSLVGSDP
3296	A	1	838	GTRGGVGPGDNGGVEAGAKPGAAAIPLRGDGS GETGPGRVAPGEVRGSPRGHVAGPEGPREVLFFF FLPSSKPASEVINEYSWKVDFLKGMLQAEKLTSS SEKALANQFLAPGRVPTTARERVPATKTVHLQS RARYTSEMRSELLGTDSAEPEMDVRKRTGVAGS QPVSEKQSAAELDLVLQRHQNLQEKLAEEMLGL ARSLKTNTLAAQSVIKKDNQTLSHSLKMADQNL EKLKTESERLEQHTQKSVNWLLWAMLIIVCFIFIS MILFIRIMPKLK
3297	A	46	617	HKQPAGFLGLWLGTETYTISFPGPETFGLGLSHA TGIPGSPACRQPVVGLHSLHNYRMAMVSAMSW VLYLWISACAMLLCHGSLQHTFQQHHLHRPEGG TCEVIAAHRCCNKNRIEERSQTVKCSCLPGKVAG TTRNRPSCVDASIVIGKWWCEMEPCLEGEECKTL PDNSGWMCATGNKIKTTRIHPRT
3298	A	157	748	IQPPDPRNMTLAAYKEKMKELPLVSLFCSCFLAD PLNKSSYKYEADTVDLNWCVISDMEVIELNKCT SGQSFEVILKPPSFDGVPEFNASLPRRRDPSLEEIQ KKLEAAEERRKYQEAELLKHLAEKREHEREVIQ KAIEENNNFIKMAKEKLAQKMESNKENREAHLA AMLERLQEKDKHAEEVRKNKELKEEASR
3299	A	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAGVA GGAPRRRTPVTMWRLLARASAPLLRVPLSDSWA LLPASAGVKTLLPVPSFEDVSIPEKPKLRFIERAPL VPKVRREPKNLSDIRGPSTEATEFTEGNFAILALG GGYLHWGHFEMMRLTINRSMDPKNMFAIWRVP APFKPITRKSVGHRMGGGKGAIDHYVTPVKAGR LVVEMGGRCEFEEVQGFLDQVAHKLPFAAKAVS RGTLEKMRKDQEERERNNQNPWTFERIATANML GIRKVLSPYDLTHKGKYWGKFYMPKRV
3300	A	2	1847	FVAGGPRGSGSAAETMPEIRVTPLGAGQDVGRS CILVSIAGKNVMLDCGMHMGFNDDRRFPDFSYI TQNGRLTDFLDCVIISHFHLDHCGALPYFSEMVG YDGPIYMTHPTQAICPILLEDYRKIAVDKKGEAN FFTSQMIKDCMKKVVAVHLHQTVQVDDELEIKA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YYAGHVLGAAMFQIKVGSESVVYTGDYNMTPDRHLGAAWIDKCRPNLLITESTYATTIRDSKRCRERDFLKKVHETVERGGKVLIPVFALGRAQELCILLETFWERMNLKVPIYFSTGLTEKANHYYKLFIPWTNQKIRKTFVQRNMFEFKHIKAFDRAFADNPGPMVVFATPGMLHAGQSLQIFRKWAGNEKNMVIMPGYCVQGTVGHKILSGQRKLEMEGRQVLEVKMQVEYMSFSAHADAKGIMQLVGQAEPESVLLVHGEAKKMEFLKQKIEQELRVNCYMPANGETVTLPTSPSIPVGISLGLLKREMAQGLLPEAKKPRLLHGTLIMKDSNFRLVSSEQALKELGLAEHQLRFTCRVHLHDTRKEQETALRVYSHLKSVLKDHCVQHLPDGSVTVESVLLQAAAPSEDPGTKVLLVSWTYQDEELGSFLTSLLKKGLPQAPS
3301	A	2	349	CIRTEPAAAFRRLGALSGAAALGFASYGAHGAQFPDAYGKELFDKANKHHFLHSLALLGVPHCRKPLWAGLLLASGTTLFCTSFYYQALSGDPSIQTLAPAGGTLLLLGWLALAL
3302	A	59	1184	LRRNCSALGGLFQTIISDMKGSYPVWEDFINKAGKLQSQLRTTVVAAAAFLDAFQKVADMATNTRGGTREIGSALTRMCMRHRSIEAKLRQFSSALIDCLINPLQEQMEEWKKVANQLDKDHAKEYKKARQEIKKKSSDTLKLQKKAKKGRGDIQPQLDSALQDVNDKYLLLEETEKQAVRKALIEERGRFCTFISMLRPVIEEEISMLGEITHLQTISEDLKSLTMDPHKLPSSSEQVILDLKGSDYSWSYQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSSGSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSHDSGFISQDAFQSKSPSPMPPEAPNQRRKEKREPDPNGGGPTTASGPPAAAEEAQRPRSM
3303	A	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKSTRGHSSLLPPSQDFVAGLSVILRGTVDDRLNWAFNLYDLNKDGCITKEEMLDIMKSIYDMMGKYTYPALREEAPREHVESFFQKMDRNKDGVVTIEEFIESCQKDENIMRSMQLFDNVI
3304	A	40	432	ISEAASGAFQAR*FYQM\LEQKTDALGKQSVNRGFTKDKTLSSIFNIEMVKEKTAEEIKQIWQQYFAAKDTVYAVIPAEKFDLIWNRAQSCPTFLCALPRREGYEFFVGQWTGTELHFHCTYKYSDPEGKA
3305	A	2	483	LDACSTGPYSRSTHASADAWADAWVVVVLKVVGMTLFLLYFPQIFNKSNDGFTTTRSYGTVSQIFGSRSPSPNGFITTRSYGTVCPKDWEFYQARCFFLIHL*\SSWNESWDFCKGKGCTLAIVDNSETLKLLHDLHDAEKNYIALPYRSSKYMSTCNGTF
3306	A	2	872	TLSSACLIGDAWKELTIVAGAVSNQLLVWYPATALADNKPVAPDRRISGHVGIIFSMSYLESKGLLATASEDRSVRIWKGGDLRVPGGRVQNIGHCFGHSARVWQVKLLENYLISAGEDCVCLVWSHEGEILQAFRGHQGRGIRAIAAHERQAWVITGGDDSGIRLWHLVGRGYRGLG/DLGSLLQVP**ARYTQGCDSGWLLATAGSD*YRGPVSL*RRGQVLGAAARG*TFPVLLPAGGSSWSRGLRIVCYGQWGRSCQGCPHQHSNCCCGPDPVSWEGAQLELGPAWL
3307	A	2	927	RTSRVEKGLRKAGAAVTMESDEWFSQALPANTSAQKAELIALTQAIRWGKDINVNTDSRYAFATVH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VRGAICQERRLLTSAEKAIKNKNPPSSKPNRSSS\FWGTTCDQVNAKQGPKPSPGHRLRRNLPGEKWEIDFTKVKPHQAGYKYLLVLVDTFSGWTEAFATKNETVNMVVKFLLNEIIPRHGLPVAIGSDNGPAFALSIV*SVSKALNIQWKLHCAYRPQSSGQVERMNCTLKNTLTKLILETGVNWVSLLPLALLRVRCTPYWAGFLPFEIMYGRVLPILPKLRDAQLAKISQTNLLQYLQSP
3308	A	490	1077	NSPSLDFNDNEDIPTELSDSSDTHDEGEVQAFYEDLSGRQYVNEVFNFSVDKLYDLLFTNSPFQRDFMEQRRFSDIIFHPWKKEENGNQSRVIPYTITLTNPLEHKTATVRETQTMYKASQESECYVIDAEVLTHDVPYHDYFYTINRYTLTRVARNKSRLRVSTELRYRKQPWGLVKTFIEKNFWSGLEDYFRHL
3309	A	490	1077	NSPSLDFNDNEDIPTELSDSSDTHDEGEVQAFYEDLSGRQYVNEVFNFSVDKLYDLLFTNSPFQRDFMEQRRFSDIIFHPWKKEENGNQSRVIPYTITLTNPLEHKTATVRETQTMYKASQESECYVIDAEVLTHDVPYHDYFYTINRYTLTRVARNKSRLRVSTELRYRKQPWGLVKTFIEKNFWSGLEDYFRHL
3310	A	2	1198	SPLCHPGLSRER/S*SEAKLRSGRYC*KRQVEAPL*RPGL*TMAASDTERDGLAPEKTSPDRDKKKEQSEVSVSPRASKHHYSRSRSRSRERKRKSDNEGRKHRSRSRSKEGRRHESKDKSSKKHKSEEHNDKEHSSDKGRERLNSSENGEDRHKRKERKSSRGRSHSRSRSRERRHRSRSRERKKSRSRSRERKKSRSRSRERKKSRSRSRERKRRIRSRSRSRSRHRHRTRSRSRTRSRSRDRKKRIEKPRRFSRSLSRTPSPPPFRGRNTAMDAQEALARRLERAKKLQEQREKEMVEKQKQQEIAAAAAATGGSVLNVAALLASGTQVTPQIAMAAQMAALQAKALAETGIAVPSYYNPAAVNPMKFAEQEKKRKMLWQGKKEGDKSQSAGNMGKN
3311	A	177	4	PIQIPPRITPPRPSPHLLTPRTGSSPPPPRAPSPPHPTPGPAHDFPPLSAVLSGHTKT
3312	A	3	426	LESPRH*PPCWGPLIWALTVSSVPSPTPELSCILKSP/RPACPV/PGLWPSLLSPAPPQSSGPLLGLSPCPGAGQWPSPLSPAPPPSSDPLSGLSPCPGAGPRSSP\SASAPCRAVPLSPRRLTWPPHLQVGILIPTGRPWKNL
3313	A	162	2	QLQNLASRGCL*SQLLRRLRRENRLNPGGGGCSEIAP\CTPAWVTQRDFFRKKK
3314	A	162	2	QLQNLASRGCL*SQLLRRLRRENRLNPGGGGCSEIAP\CTPAWVTQRDFFRKKK
3315	A	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWKGRKHASRTARAHVFHPIRQSIRSPVRGRPGDPRAAHTRSAGTRLQCKASRGG*GKGPAPTR*EGGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPGAAAAQP**TPRCPAALRAGAHIGRVGRPY
3316	A	3	2307	NHLGTLMQNWDSSSRVPFSSGQHSTQSFPPSLMSKSNSMLQKPT\AYVRPMDGQESMEPKLSSEHYSSQSHGNSMTELKPSSKAHLTKLKIPSQPLDASASGDVSCVDEILKEMTHSWPPPLTAIHTPCKTEPSKFPFPTKESQQSNFGTGEQKRYNPSKTSNGHQSKSMLKDDLKLSSSEDSDGEQDCDKTMPRSTPGSNSEPSHHNSEGADNSRDDSSSHSGSESSSGSDSESESSS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SDSEANEPSQSASPEPEPPPTNKWQLDNWLNKV NPHKVSPASSVDSNIPSSQGYKKEGREQGTGNSY TDTSGPKETSSATPGR\APKPIQKGSESGRGRQKS PAQSDSTTQRRTVGKKQPKKAEKAAAEEPRGGL KIESETPVDLASSMPSSRHKAATKGSRKPNIKKES KSSPRPTAEKKKYKSTSKSSQKSREIIETDTSSSDS DESESLPPSSQTPKYPESNRTPVKPSSVEEEDSFFR QRMFSPMEEKELLSPLSEPDDRYPLIVKIDLNLLT RIPGKPYKETEPPKGEKKNVPEKHTREAQKQASE KVSNKGKRKHKNEDDNRASESKKPKTEDKNSA GHKPSSNRESSKQSAAKEKDLLPSPAGPVPSKDP KTEHGSRKRTISQSSSLKSSSNSNKETSGSSKNSS STSKQKKTEGKTSSSSKEVKVKAPSSSSNCPPSAP TLDSSKPRRTKLVFDDRNYSADHYLQEAKKLKH NADALSDRFEKAVYYLDAVVSFIECGNALEKNA QESKSPFPMYSETVDLI
3317	A	496	2	NLLQDEKLVHSYPYDWRTQETCGYIVPARQWFI N\TRDIKTAAKELLKKVKFIPGSALNGMVEMMD RRPYWCISRQRVWGVPIPVFHHKTKDEYLINSQT TEHIVKLVEQHGSDIWWTLPPEQLLPKEVLSEVG GPDALEYVPGQDILDIWFDSGTSWSYVLPGPD
3318	A	2	512	AWHEGDSRSDQCHHPYNYGFDYYYGMPFTLVD SCWPDPSRNTELAFESQLWLCVQLVAIAILTLTF GKLSGWVSVPWLLIFSMILFIFLLGYAWFSSHTSP LYWDCLLMRGHEITEQPMKAE\RAGSIMVKEAIF LFRKGHSKGKLFLLFFLPFLQVHKTFPTTDGFHW AP
3319	A	407	1	SSLHRSPRPASPLPVPEAP\SFLPVPAPKPSALPPFS LSGAPSSASTFSPHSSPSPASPTPAPSPQSPFPSRPT SPPSLTPTRRPPLPADRRGPHLLYQPLHAPLEAAA TGPE/PSAAAGRLPRPRPPWRAAYPASR
3320	A	4037	3432	QMSEAVAEKMLQYRRDTAGWKICREGNGVSVS WRPSVEFPGNLYRGEGIVYGTLEEVWDCVKPAV GGLRVKWDENVTGFEIIQSITDTLCVSRTSTPSAA MKLISPRDFVDLVLVKRYEDGTISSNATHVEHPL CPPKPGFVRGFNHPCGCFCEPLPGEPTKTNLVTFF HTDLSGYLPQNVVDSFFPRSMTRFYANLQKAVK
3321	A	37	360	SHSASGAGRPAAPAADLRPAPNGQRPGPRLGAR ALWLPPRGRPDEAGRLPGEHLPQVPWDPGLTRS PSPRGPCRGAARAGHVGETPAPWGCPPPCAWEH KGPGSEGTP
3322	A	1	420	AIVEDKHSGRSYDITSDLGNVLTSTSIAKTVNG*A ESSDSGAESDEEDAQEDLMGAYHSDIDKKMMKI VADHKNLEVIVTNGYDKDGFVHDIQNDIHASSSL NGRSTVHVKPIDENLGQTGKSAVCIHQDINDDH VEDVT
3323	A	8	459	DTLSLNCTLPETLPMTPSF*LSFL*FPGLARAKSIP TKTYSNEVVTLWYRPPDILLGSTDYSTQIDMW*G QVEVWQGPCGKGGGLVTTATQPAAFLFTVPSLP RGVGCIFYEMATGRPLFPGSTVEEQLHFIFRILSE EAWALCAVETHR
3324	A	1276	466	PGSTHASARITIY*L*IILSNATEVDNNFSKPPPFFP AGAPPASSSSSSSSSPPTVSTAPPLIPPPGFPPPPG APPPSLIPTIESGHSSGYDSRSARAFPYGNVAFPH LPGSAPSWPSLVDTSKQWDYYARSSSSSSSSSSS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SSSPRDRDRER*RTRERERERDHSPTPSVFNSDEE RYRYREYAERGYERHRASREKEERHRERRHREK EETRHKSSRSNSRRRHESEEGDSHRRHKHKKSKR SKEGKEAGSEPAPEQESTEATPAE
3325	A	266	3312	TCLFSASCSSLPSPSSSFALLSTENTQRTYRVNPD GSLRVTFASGMEIGLSSEPHILAGAVNPTLGKCNI SLPGEHNANLISVL**GEQGCA*NVFHISFS*AHN RNLLSIDFDHITRTGKIYDDHRKFTLRILYDQTGR PILWSPVSRYNEVNITYSPSGLVTFIQRGTWNEK MEYDQSFL*SPQL*LSIICYSAFVSFQSVMLLLHS QRRYIFEYDQPDCLLSVTMPSMVRHSLQTMLSV GYYRNIYTPPDSSTSFIQDYSRDGRLLQTLHLGTG RRVLYKYTKQARLSEVLYDTTQVTLTYEESSGD LSDSSTLIA*LLTVFVLVPAGPLIGRQIFRFSEEGL VNARFDYSYNNFRVTSMQAVINETPLPIDLYRYV DVSGRTEQFGKFSVINYDLNQVITTTVMKHTKIF SANGQVIEVQYEILKAIAYWMTIQYDNVGRMVI CDIRVGVDANITRYFYEYDADGQLQTVSVNDKT QWRYSYDLNGNINLLSHGKSARLTPLRYDLRDRI TRLGEIQYKMDEDGFLRQRGNDIFEYNSNGLLQ KAYNKASGWTVQYYYDGLGRRVASKSSLGQHL QFFYADLTNPIRVTHLYNHTSSEITSLYYDLQGH LIAMELSSGEEYYVACDNTGTPLAVFSSRGQVIK EILYTPYGDIYHDTYPDFQVIIGFHGGLYDFLTKL VHLGQRDYDVVAGRWTTPNHHIWKQLNLLPKP FNLSTKLIKYGIFHFLFLILCLTDIRSWLELFGFQL HNVLPGFPKPELENSPSI*QMSNSMLHLLCASLS* TILGIQCELQKQLRNFISLDQLPMTPRYNDGRCLE GGKQPRFAAVPSVFGKGIKFAIKDGIVTADIIGVA NEDSRRLAAILNNAHYLENLHFTIEGRDTHYFIK LGSLEEDLVLIGNTGGRRILENGVNVTVSQMTSV LNGRTRRFADIQLQHGALCFNIRYGTTVEEEKNH VLEIARQRAVAQAWTKEQRRLQEGEEGIRAWTE GEKQQLLSTGRVQGYDGYFVLSVEQ
3326	A	290	1041	KACLHLLSSFLTSNFLFNPLLPDSLYSVEARSQRA NLGPCRRKRLQTLMRLAAGFQYSSHKDPSLSAK EKHTDYHNEARGPWPGWVG*RTADGSCGRGPD GAHHPGPKSSSWRASRLLPGLGGSHHLDAYVGR DLECGTPAPLQLEIPPQPRGHPAPIPTGQAGPRDS GPGASP*VETRPLTDGRR*PGVRPVGWTPAHPAG TLRPRGAVEPSVSACGKWAPSPTSQGCCEGRCD AVPKHRAWRTPLCSQ
3327	A	1	418	CSECGKSFCKKSKFIIHQRTHTGEKPYECNQCGK SFCQKGTLTVHQRTHTGEKPYECNECGKNFYQK LHLIQHQRTHSGEKPYECSYCGKSFCQKTHLTQH QRTHSGERPYVCHDCGKTFSQKSALNDHQKIHT GVKLY
3328	A	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENELDED MHQKIAREMNLSETAFIRKLHPTDNFAQRSCFGL IWFTPTTDLQILTSSILPSIL
3329	A	45	419	EELSCWQIWQQIANDLTRCQDSMINNSQCHKQG DFPYQVGTELSIQISEDENYIVNKADGPNNTGNP EFPILRTQDSWRKTFLTESQRLNRDQQISIKNKLC QCKKGVDPIGWISHHDGHRVHKR
3330	A	64	430	FWRNFTGLAPAAAVATTTSSSTMRFTSISNSLTST

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AAIGLSFTTSTTTTATFTTNTTTTITSGFTVNQNQ LLSRGFENLVPYTSTVSVVTTPVMTYGHLEGLIN EGNLELEIKRRLSSQATQ
3331	A	3	407	TFGCSCTDCFFQKCCPAEAGVLLAYNKNQQIKIP PGTPIYECNSRCQCGPDCPNRIVQKGTQYSLCIFR TSNGRGWGVKTLVKIKRMSFVMEYVGEVITSEE AERRGQFYDNKGITYLFDLDYESDEFTVDAARY
3332	A	25	461	PAADFVLQARPTRADILGIHSKYDEVRKAGACFY KMTGLGPGPQALYNGEPFKHEEMNIKELKMAVL QRMMDASVYLQREVFLGTLNDRTNAIDFLMDR NNVVPRINTLILRTNQQYLNLLSTSVTADAEDFS TFFFLDSQDKSA
3333	A	317	54	AWIIFLPPLTSCPLWAPGTKHKTILEARSGLGPIK AYPRLGPPTPGEPEAPAQDRTFHCEICNVKVNSK VQLKQHISSRRHEIVDPV
3334	A	304	410	AGPSLPSNLRQIFQSLPPFMDILLLLLFFMIIFAI
3335	A	19	418	VESRNSRVQPRVRLNDRTNAIDFLMDRNNVVPRI NTLILRTNQQYLNLISTSVTADVEDFSTFFFLDSQ DKSAVIAKNMYYLTQDDESIISAATLWIIADFDK PSGRKLLFNALKHMITSVHSRVGIIYNPFF
3336	A	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLLNR VLERLAGGATRDSAASDILLDDIVLTHSLFLPTEK FLQELHQYFVRAGGMEGPEGLGRKQACLAMLL HFLDTYQGLLQEEEGAGHIIKDLYLLIMKDESLY QGLREDTLRLHQLVETVELKIPEENQPPSKQVKP LFRHFRRIDSCLQTRVAFRGSDEIFCRVYMPDHS YVTIRSRLSASVQDILGSVTEKLQYSEEPAGREDS LILVAVSSSGEKVLLQPTEDCVFTALGINSHLFAC TRDSYEALVPLPEEIQVSPGDTEIHRVEPEDVANH LTAFHWELFRCVHELEFVDYVFHGE
3337	A	444	43	KILLCLANQFPDISFCPALPAVVALLLHYSIDEAE CFEKACRILACNDPGRRLIDQSFLAFESSCMTFGD LVNKYCQAAHKLMVAVSEDVLQVYADWQRWL FGELPLCYFARVFDVFLVEGYKVLYRVALAXXF
3338	A	1	398	FRGKVRGRSAEMPGSDTALTVDRTYSDPGRHHR CKSRVERHDMNTLSLPLNIRRGGSDTNLNFDVPD GILDFHKVKLTADSLKQKILKVTEQIKIEQTSRDG NVAEYLKLVNNADKQQAGRIKQVFEKKNQK
3339	A	1	665	AAAASNWGLITNIVNSIVGVSVLTMPFCFKQCGI VLGALLLVFCSWMTHQSCMFLVKSASLSKRRTY AGLAFHAYGKAGKMLVETSMIGLMLGTCIAFYV VIGDLGSNFFARLFGFQVGGTFRMFLLFAVSLCI VLPLSLQRNMMASIQSFSAMALLFYTVFMFVIVL SSLKHGLFSGQWLRRVSYVRWEGVFRCIPIFGMS FACQSQVLPTYDSLDEPSV
3340	A	198	367	LLPLQVLQEAFSRCVAVLTRSSKPSDMSVQVCG YISKCYSVAAQFEECREKITEMP
3341	A	562	277	HSVIKRTPRKYLAEIVLIDDFSNKEHLKEKLDEYI KLWNGLVKVFRNERREGLIQARSIGAQKAKLGQ VLIYLDAHCEVAVNWYAPLVAPISKDR
3342	A	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMWWE SLLLLTAYFCYVVFMKFNVQVEKWVKQMINRN KVVKVTAPEAQAKPSAARDKDEPTLPAKPRLQR GGSSASLHNSLMRNSIFQNKIHTLDPHV
3343	A	1	385	FRVDNSEEWKDVFIISSERSFKLDSLKCGTWYKV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KLAAKNSVGSGRISEIIEAKTHGREPSFSKDQHLF THINSTHARLNLQGWNNGGCPITAIVLEYRPKGT WAWQGLRANSSGEVFLTELREATWY
3344	A	351	147	SPACITSSLSQHIADPRAAPTEVKVRVMNSTAISL QWNRVYSDTVQGQLREYRVRKPAPDSPNYPAH
3345	A	351	147	SPACITSSLSQHIADPRAAPTEVKVRVMNSTAISL QWNRVYSDTVQGQLREYRVRKPAPDSPNYPAH
3346	A	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLKMP RGIAIDWVAGNVYWTDSGRDVIEVAQMKGENR KTLISGMIDEPHAIVVDPLRGTMYWSDWGNHPK IETAAMDGTLRETLVQDNIQWPTGLAVDYHNER LYWADAKLSVIGSIRLNGTDPIVAADSKRGLSHP FSIDVFEDYIYGVTYINNRVFKIHKFGHSPLVNLT GGLSHASDVVLYHQHKQPEVTNPCDRKKCEWL CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPPPD APRPGTCNLQCFNGGSCFLNARRQPKCRCQPRY TGDKCELDQCWEHCRNGGTCAASPSGMPTCRCP TGFTGPKCTQQVCAGYCANNSTCTVNQGNQPQ CRCLPGFLGDRCQYRQCSGYCENFGTCQMAAD GSRQCRCTAYFEGSRCEVNKCSRCLEGACVVNK QSGDVTCNCTDGRVAPSCLTCVGHCSNGGSCTM NSKMMPECQCPPHMTGPRCEEHVFSQQQPGHIA SILIP
3347	A	974	666	SPEMESHPITQAGVQWHHLSSLQPLPPGFK*FSCF SLPE*LGYRHVPPCLANSVFSVEMG\FLHVGQAG LELLTSGDLPALASQSAGITG\SHRARPENGFENIF
3348	A	1	1171	LSKITMPVICNEPLSFIQRLTEYM*HTYFIHRPSSL SDPVDRMQCVAAFAVSAVASQWERTGKPFNPLL GETYELVRDDLGFRLISEQVSHHPPISAFHAEGLN NDFIFHGSIYPKLKFWGKSVEAEPKGTITLELLEH NEAYTWTNPTCCVHNIIVGKLWIEQYGNVEIINH KTGDKCVLNFKPCGLFGKELHKVEGYIQDKSKK KLCALYGKWTECLYSVDPATFDAYKKNDKKNT EEKKNSKQMSTSEELDEMPVPDSESVFIIPGSVLL WRIAPRPPNSAQMYNFTSFAMVLNEVDKDMESV IPKTDCRLRPDIRAMENGEIDQASEEKKRLEEKQ RAARKNRSKSEEDWKTRWFHQGPNPYNGAQD WIYSGSYWDRNYFNLPDIY
3349	A	403	497	NFASSSGKYLRTQKIKCLNNKFTPFPTTEKK*SQS VRPP*SNRIY*ILQS*NISFS*LPN*NFASSSGKYLR TQKIKCLNNKFTPFPTTEKK
3350	A	1	712	GAPAQDCICLPFPFHSSFLESDIRKPARRKIQTTNP DFLLLLFMSVPVVSAPPFCPPAEGSRDGRPKASV ARPAAVHEHHSPRDCGHLPDVIRSSLGGWQPH*P AQPENRLL*LLPVE*GHQHPTVSPVP*AGSPGGAS GWPGPGQAWRVRVPGPHPLCPPASPPSPVQQ**E SVAAGSGLPGCVLCAAGRRPGPLPLLCVEVGQA LPPGAWVSSSGQRPGLTHPLAYSHGCVPSEG
3351	A	1	428	MAAVVAATALKGRGARNARVLRGILAGATANK ASHNRTRALQSHSSPEGKEEPEPLSPELEYIPRKR GKNPMKAVGLAWAIGFPCGILLFILTKREVDKDR VKQMKARQNMRLSNTGEYESQRFRASSQSAPSP DVGSGVQT
3352	A	2	841	RTLFRGRRRREDDRISRPHPSTAESKAPTPKFDLL ASNFPPLPGSSSRMPGELVLENRMSDVVKGVYK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EKDNEELTISCPVPADEQTECTSAQQLNMSTSSP CAAELTALSTTQQEKDLIEDSSVQKDGLNQTTIP VSPPSTTKPSRASTASPCNNNINAATAVALQEPR KLSYAEVCQKPPKEPSSVLVQPLRELRSNVVSPT KNEDNGAPENSVEKPHEKPEARASKDYSGFRGN IIPRGAAGKIREQRRQFSHRAIPQGVTRRNGKEQ YVPPRSPK
3353	A	1054	587	IATPTWTAPLTATPTPAHQYGPARVPNGAPRLEP PPGKRECRVGQYVVDLTSFEQLALPVLRNADCS SGPGQRVCVIDEIGKMELFSQLFIQAVRQTLSTPG TIILGTIPVPKGKPLALVEEIRNRKDVKVFNVTKE NRNHLLPDIVTCVQSSRK
3354	A	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVVER VLTFLPAKALLRVACVCRLWRECVRRVLRTHRS VTWISAGLAEAGHLEGHCLVRVVAEELENVRILP HTVLYMADSETFISLEECRGHKRARKRTSMETA LALEKLFPKQCQVLGIVTPGIVVTPMGSGSNRPQ EIEIGESGFALLFPQIEGIKIQPFHFIKDPKNLTLER HQLTEVGLLDNPELRVVLVFGYNCCKVGASNYL QQVVSTFSDMNIILAGGQVDNLSSLTSEKNPLDI DASGVVGLSFSGHRIQSATVLLNEDVSDEKTAEA AMQRLKAANIPEHNTIGFMFACVGRGFQYYRAK GNVEADAFRKFFPSVPLFGFFGNGEIGCDRIVTG NFILRKCNEVKDDDLFHSYTTIMALIHLGSSK
3355	A	1	707	GTSSGLGGDRLAAPGPSPPSFYPQGRGERAYDIY SRLLRERIVCVMGPIDDSVASLVIAQLLFLQSESN KKPIHMYINSPGGVVTAGLAIYDTMQYILNPICT WCVGQAASMGSLLLAAGTPGMRHSLPNSRIMIH QPSGGARGQATDIAIQAEEIMKLKKQLYNIYAKH TKQSLQVIESAMERDRYMSPMEAQEFGILDKVL VHPPQDGEDEPTLVQKEPVEAAPAAEPVPAST
3356	A	352	338	FNYNFCRNLHMPSFLV*PGMCGLLAKHLSFHIVG AFLIT/LGVAALCKFAVA*PRKKAYADFYRNYN* IKEFEVRKANISQSTK
3357	A	1	403	ALGSCGGLLGTGLLKGTMSGTLWSKGIFAGYKR RIRIQREHTAVLKIEG\VYARDETEFYLRMICANV YKANNNTVTPVLTPDKTRVMWRKVTQAHGISI MVRAQFRTNLPADAIGHRIRMML*PSRMYTTEPS
3358	A	71	2897	FCSKDKCCLYLPDSINRSKSCTAKPGAHSQDRHA VMDSERQVKDTDDIESPKRSIRDSGYIDCWDSER SDSLSPPRHGRDDSFDSLDSFGSRSRQTPSPDVVL RGSSDGRGSDSESDLPHRKLPDVKKDDMSARRT SHGEPKSAVPFNQYLPNKSNQTAYVPAPLRKKK AEREEYRKSWSTATSPAGLGKKALQDYGPRT\PV S\DDAESTSMFDMRCEEEAAVQPHSRARQEQLQ LINNQLREEDDKWQDDLARWKSRKRSVSQDLIK KEEERKKMEKLLAGEDGTSERRKSIKTYREIVQE KERRERELHEAYKNARSQEEAEGILQQYIERFTIS EAVLERLEMPKILERSHSTEPNLSSFLNDPNPMK YLRQQSLPPPKFTATVETTIARASVLDTSMSAGS GSPSKTVTPKAVPMLTPKPYSQPKNSQDVLKTFK VDGKVSVNGETVHREEEKERECPTVAPAHSLTK SQMFEGVARVHGSPLELKQDNGSIEINIKKPNSV PQELAATTEKTEPNSQEDKNDGGKSRKGNIELAS SEPQHFTTTVTRCSPTVAFVEFPSSPQLKNDVSEE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KDQKKPENEMSGKVELVLSQKVVKPKSPEPEAT LTFPFLDKMPEANQLHLPNLNSQVDSPSSEKSPV TTPFKFWAWDPEEERRRQEKWQQEQERLLQER YQ\KEQDK\LKEE\WEKAQKEVEEEERRYYEEEP* II\EDPVVPFTVSSSSADQLSTSSSMTEGSGTMNKI DLGNCQDEKQDRRWKKSFQGDDSDLLLKTRES DRLEEKGSLTEGALAHSGNPVSKGVHEDHQLDT EAGAPHCGTNPQLAQDPSQNQQTSNPTHSSEDV KPKTLPLDKSINHQIESPSERRKSISGKKLCSSCGL PLGKGAAMIIETLNLYFHIQCFRCG\ICKGQLGDA VSGTDVRIRNGLLNCNDCYMRSRSAGQPTTL
3359	A	3	368	EVTASREGRGACAWECGSSRGPWGLLRGTFAPV RAATP*S*LPKGSLRHRP*/CPPPVHLPPKSSCPPR AWAGRATSM*TSSYSSEYQPQTP*ALVTLPPRSY YLLTHLLTLTHLHHQILFEP
3360	A	2	392	ARGIGSLGRDHSGSGGGTGMAGAWVRKAADYV RSKDFRDYLMSTHFWGPVANWGLPIAAITDMK\ KSPEIISRRMTFAL*CYSLTFVRFAHYVQ\PWNWL MLGCHTAVDFDQLISSMPCISHGMTASASAL
3361	A	4619	532	LLLGRANSPPYNSVVRTLPPATLLLRRAGWESF WSCQSRSPWPPRPEVRAPAKGPRGVAGAAGACS AGARLGDAAGGDPASGQAARGCGARAPRGLGR TARARDTAMEDAGAAGPGPEPEPEPEPEPEPAPE PEPEPKPGAGTSEAFSRLWTDVMGILDGSLGNID DLAQQYADYYNTCFSDVCERMEELRKRRVSQD LEVEKPDASPTSLQLRSQIEESLGFCSAVSTPEVE RKNPLHKSNSEDSSVGKGDWKKKNKYFWQNFR KNQKGIMRQTSKGEDVGYVASEITMSDEERIQL MMMVKEKMITIEEALARLKEYEAQHRQSAALDP ADWPDGSYPTFDGSSNCNSREQSDDETEESVKF KRLHKLVNSTRRVRKKLIRVEEMKKP\STEGGEE HVFENSPVLDERSALYSGVHKKPLFFDGSPEKPP EDDSDSLTTSPSSSSLDTWGAGRKLVKTFSKGES RGLIKPPKKMGTFFSYPEEEKAQKVSRSLTEGEM KKGLGSLSHGRTCSFGGFDLTNRSLHVGSNNSDP MGKEGDFVYKEVIKSPTASRISLGKKVKSVKET MRKRMSKKYSSSVSEQDSGLDGMPGSPPPSQPD PEHLDKPKLKAGGSVESLRSSLSGQSSMSGQTVS TTDSSTSNRESVKSEDGDDEEPPYRGPFCGRARV HTDFTPSPYDTDSLKLKKGDIIDIISKPPMGTWMG LLNNKVGTFNFIYVDVLSED\EEKPKRPTRRRRK GRPPQPKSVEDLLDRINLKEHMPTFLFNGYEDLD TFKLLEEEDLDELNIRDPEHRADLLTAVELLQEY DSNSDQSGSQEKLLVDSQGLSGCSPRDS*CYESS ENLENGKTRKASLLSAKSSTEPSLKAFSRNQLGN YPTLPLMKSGDALKQGQEEGRLGGGLAP\DTSKS CDPPGC*LVLN\KNRRKPPSFPSCRSC\ETL\EGPQ TVDTWPRSHSLDDLQVEPGAEQDVPTEVTEPPPQ IVPEVPQKTTASSTKAQPLEQDSAVDNALLLTQS KRFSEPQKLTTKKLEGSIAASGRGLSPPQCLPRNY DAQPPGAKHGLARTPLEGHRKGHEFEGTHHPLG TKEGVDAEQRMQPKIPSQPPPVPAKKSRERLANG LHPVPMGPSGALPSPDAPCLPVKRGSPASPTSPSD CPPALAPRPLSGQALGSPPSTRPPPWLSELPENTS LQEHGVKLGPALTR\KVSCARGVDLETLTENKL\

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HAEGIRSSRREPYS*LRHGRCGI\P\EALVQRYAED LDQPERDVAANMDQIRVKQLRKQHRMAIPSGGL TEICRKPVSPGCIS\SVSDWLISIGLPMYAGTLSTA GFSTL\SQVPSLSHTCLQEAG\ITEERHIRK\LLSAA RLFKLPPGPEAM
3362	A	1	4653	FRGGVGYAHTLHLLPFAGSSVVLARARRTDRWT SGLVEMATLSLTVNSGDPPLGALLAVEHVKDDV SISVEEGKENILHVSENVIFTDVNSILRYLARVAT TAGLYGSNLMEHTEIDHWLEFSATKLSSCDSFTS TINELNHCLSLRTYLVGNSLSLADLCVWATLKG NAAWQEQLKQKKAPVHVKRWFGFLEAQQAFQS VGTKWDVSTTKARVAPEKKQDVGKFVELPGAE MGKVTVRFPPEASGYLHIGHAKAALLNQHYQV NFKGKLIMRFDDTNPEKEKEDFEKVILEDVAML HIKPDQFTYTSDHFETIMKYAEKLIQEGKAYVDD TPGEQIKAEREQRIESKHRKNPIEKNLQMWEEMK KGSQFGHSCCLRAKIDMSSNNGCMRDPTLYRCK IQPHPRTGN*Y\NV\YPTYDFACPIVDSIEGVTHAL RTTEYHDRDEQFYWIIEALGIRKPYIWEYSRLNL NNTVLSKRKLTWFVNEGLVDGWDDPRFPTVRG VLRRGMTVEGLKQFIAAQGSSRSVVNMEWDKI WAFNKKVIDPVAPRYVALLKKEVIPVNVPEAQE EMKEVAKHPKNPEVGLKPVWYSPKVFIEGADAE TFSEGEMVTFINWGNLNITKIHKNADGKIISLDAK LNLENKDYKKTTKVTWLAETTHALPIPVICVTYE HLITKPVLGKDEDFKQYVNKNSKHEELMLGDPC LKDLKKGDIIQLQRRGFFICDQPYEPVSPYSCKEA PCVLIYIPDGHTKEMPTSGSKEKTKVEATKNETS APFKERPTPSLNNNCTTSEDSLVLYNRVAVQGD VVRELKAKKAPKEDVDAAVKQLLSLKAEYKEK TGQEYKPGNPPAEIGQNISSNSSASILESKSLYDE VAAQGEVVRKLKAEKSPKAKINEAVECLLSLKA QYKEKTGKEYIPGQPPLSQSSDSSPTRNSEPAGLE TPEAKVLFDKVASQGEVVRKLKTEKAPKDQVDI AVQELLQLKAQYKSLIGVEYKPVSATGAEDKDK KKKEKENKSEKQNKPQKQNDGQRKDPSKNQGG GLSSSGAGEGQGPKKQTRLGLEAKK\EENLADW YSQVITKSEMIEYHDISGCYILRPWAYAIWEAIKD FFDAEIKKLGVENCYFPMFVSQSALEKEKTHVA DFAPEVAWVTRSGKTELAEPIAIRPTSETVMYPA YAKWVQSHRDLPIKLNQWCNVVRWEFKHPQPF LRTREFLWQEGHSAFATMEEAAEEVLQILDLYA QVYEELLAIPVVKGRKTEKEKFAGGDYTTTIEAF ISASGRAIQGGTSHHLGQNFSKMFEIVFEDPKIPG EKQFAYQNSWGLTTRTIGVMTMVHGDNMGLVL PPRVACVQVVIIPCGITNALSEEDKEALIAKCNDY RRRLLSVNIRVRADLRDNYSPGWKFNHWELKG VPIRLEVGPRDMKSCQFVAVRRDTGEKLTVAEN EAETKLQAILEDIQVTLFTRASEDLKTHMVVANT MEDFQKILDSGKIVQIPFCGEIDCEDWIKKTTARD QDLEPGAPSMGAKSLCIPFKPLCELQPGAKCVCG KNPAKYYTLFGRSY
3363	A	3797	1514	LGGAAPETMPFPVTTQGSQQTQPPQKHYGITSPIS LAAPKETDCVLTQK\LI\ETLKPFGGFLKKEEGTA SRRNFNFGKN*INLVKEWIRRNQ*KAKNLPQSVI\

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ENV\GGKIFT/FLGSYRL/GEVHTKGADIDGVCVF APRHVDRSDFFT\SFYDKLKLQEEVKDLRAVEEA FVPVIKLCFDGIEIDILFARLALQTIPEDLDLRDDS LLKNLDIRCIRSLNGCRVTDEILHLVPNIDNFRLT LRAIKLWAKRHNIYSNILGFLGGVSWAMLVART CQLYPNAIASTLVHKFFLVFSKWEWPNPVLLKQP EECNLNLPVWDPRVNPSDRYHLMPIITPAYPQQN STYNVSVSTRMVMVEEFKQGLAITDEILLSKAE WSKLFEAPNFFQKYKHYIVLLASAPTENQRLEW VGLVESKIRILVGSLEKNEFITLAHVNPQSFPAPK ENPDKEEFRTMWVIGLVFKKTENSENLSVDLTY DIQSFTDTVYRQAINSKMFEVDMKIAAMHVKRK QLHQLLPNHVLQKKKKHSTEGVKLTALNDSSLD LSMDSDNSMSVPSPTSATKTSPLNSSGSSQGRNS PAPAVTAASVTNIQATEVSVPQVNSSESSGGTSSE SIPQTATQPAISPPPKPTVSRVVSSTRLVNPPPRSS GNAATSGNAATKIPTPIVGVKRTSSPHKEESPKK TKTEEDETSEDANCLALSGHDKTEAKEQLDTETS TTQSETIQTAASLLASQKTSSTDLSDIPALPANPIP VIKNSIKLRLNR
3364	A	54	3073	SARTMSYDYHQNWGRDGGPRSSGGGYGGGPAG GHGGNRGSGGGGGGGGGGRG/WQGPASRAPER PRNRHVVREKTGAEEQ/WKRRGKREL/LVHMDE RREEQIVQLLNSVQAKNDKESEAQISWFAPEDHG YGTEVSTKNTPCSENKLDIQEKKLINQEKKMFRI RNRSYIDRDSEYLLQENEPDGTLDQKLLEDLQKK KNDLRYIEMQHFREKLPSYGMQKELVNLIDNHQ VTVISGETGCGKTTQVTQFILDNYIERGKGSACRI VCTQPRRISAISVAERVAAERAESCGSGNSTGYQI RLQSRLPRKQGSILYCTTGIILQWLQSDPYLSSVS HIVLDEIHERNLQSDVLMTVVKDLLNFRSDLKVI LMSATLNAEKFSEYFGNCPMIHIPGFTFPVVEYLL EDVIEKIRYVPEQKEHRCQFKRGFMQGHVNSQE KEEKEAIYKERWPDYVRELRRRYSASTVDVIEM MEDDKVDLNLIVALIRYIVLEEEDGAILVFLPGW DNISTLHDLLMSQVMFKSDKFLIIPLHSLMPTVN QTQVFKRTPPGVRKIVIATNIAETSITIDDVVYVID GGKIKETHFDTQNNISTMSAEWVSKANAKQRKG RAG\RVQPGSLLFICINGS*EASLLGWTIQLPEIF/R GTPLEELCLQIKVLRLGGI/GLFLSRLMDPPSNEA VLLSIRQL\RSLNALDKQEELTPLGVHLARLPVEP HIGKMILFGALFCCLDPVLTIAASLSFKDPFVIPLG KEKIADARRKELAKDTRSDHLTVVNAFEGWEEA RRRGFRYEKDYCWEYFLSSNTLQMLHNMKGQF AEHLLGAGFVSSRNPKDPESNINSDNEKIIKAVIC AGLYPKVAKIRLNLGKKRKMVKVYTKTDGLVA VHPKSVNVEQTDFHYNWLIYHLKMRTSSIYLYD CTEVSPYCLLFFGGDISIQKDNDQETIAVDEWIVF QSPARIAHLVKRAVVHMDERREEQIVQLLNSVQ AKNDKESEAQISWFAPEDHGYDKKYFFKE
3365	A	439	878	ECCNVRPLRETDLLKMKRKPRASSPVVEEQPRA NTKETRKKKSFSQPMSASTKEESQDGRRKGK*L KGRARKKNAPQKSMALRILEEGSRPTPSGHSDQL NEEL*QNELQLEQ/PEGT*LEQQSEGTQPEQQSGR MPTISTLSLSSE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3366	A	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKITRQ KHAKKHLGFFRNNFGVREPYQILLDGTFCQAAL RGRIQLREQLPRYLMGETQLCTTRCVLKELETLG KDLYGAKLIAQKCQVRNCPHFKNAVSGSECLLS MVEEGNPHHYFVATQDQNLSVKVKKKPGVPLM FIIQNTMVLDKPSPKTIAFVKAVESG\RLSQCMRK KVSNISKRNRV**KTLNRGRRKKRKKISGPNPLS CLKKKKKAPDTQSSASEKKRKRKRIRNRSNPKV LSEKQNAEGE
3367	A	40	1467	MLWGCRAKACWGPRLSDLVASLSPQRECISVHV GQAGVQIGNACWELFCLEHGIQADGTFDAQASK INDDDSFTTFFSETGNGKHVPRAVMIDLEPTVVD EVRAGTYRQLFHPEQLITGKEDAANNYARGHYT VGKESIDLVLDRIRKLTDACSGLQGFLIFHSFGGG TGSGFTSLLMERLSLDYGKKSKLEFAIYPAPQVS TAVVEPYNSILTTHTTLEHSDCAFMVDNEAIYDI CRRNLDIERPTYTNLNRLISQIVSSITASLRFDGAL NVDLTEFQTNLVPYPRIHFPLVTYAPIISAEKAYH EQLSVAEITSSCFEPNSQMVKCDPRHGKYMACC MLYRGDVVPKDVNVAIAAIKTKRTIQFVDWCPT GFKVGINYQPPTVVPGGDLAKVQRAVCMLSNTT AIAEAWARLDHKFDLMYAKRAFVHWYVGEGM EEGEFS*RPGEDLA\ALE\KDYEEVGTDSFEEENE GEEF
3368	A	3	2597	SLLEETMDEDSSLREYTVSLDSDMDDASKCLQE YDSGTGNTREALRPCPRTVSTKAQPGRSASSSG DKTTSFAEQKIRKLNHTDGESSGSSSQKTTPEGSE LNIPHAGAWAQIPEETGLPQGRDTTQLLASEMV HLMMK\LKEKR\RAI*AQKKKMEAAFTKQRQKM GRTAFLTVVKKKGDGISPLREEAAGAEDEKVYT DRAKEKESQKTDGQRSKSLADIKESMENPQAKW LKSPTTPIDPEKQGNLASPSEETLNEGEILEYTKSI EKLNSSLHFLQQEMQRLSLQQEMLMQMREQQS WVISPPQPSPQKQIRDFKPSKQAGLSSAIAPFSSD\ SPR\PTHPSSTSLLNRKSASFSVKSQRTPRPNELKI TPLNRTLTPPRSVDSLPRLRRFSPSQVPIQTRSFVC FGDDGEPQLKESKPKEEVKKEELESKGTLEQRG HNPEEKEIKPFESTVSEVLSLPVTETVCLTPNEDQ LNQPTEPPPKPVFPPTAPKNVNLIEVSLSDLKPPE KADVPVEKYDGESDKEQFDDDQKVCCGFFFKD DQKAENDMAMKRAALLEKRLRREKETQLRKQQ LEAEMEHKKEETRRKTEEERQKKEDERARREFIR QEYMRRKQLKLMEDMDTVIKPRPQVVKQKKQR PKSIHRDHIESPKTPIKGPPVSSLSLASLNTGDNES VHSGKRTPRSESVEGFLSPSRCGSRNGEKDWEN ASTTSSVASGTEYTGPKLYKEPSAKSNKHIIQNAL AHCCLAGKVNEGQKKKILEEMEKSDANNFLILF RDSGCQFRSLYTYCPETEEINKLTGIGPKSITKKM IEGLYKYNSDRKQFSHIPAKTLSASVDAITIHSHL WQTKRPVTPKKLLPTKA
3369	A	977	594	RGSGLTQEPGSVGQLALACAEGAVEWLYPAGAL RLTLGGPDPRARPGIACLRPVRPFAGAQVFAERA GGALELLLAEGPGPAGGRCVRWGPRERRALFLQ ATPHQDISRRVAAFRFELREDGRPEIAP
3370	A	345	1383	DLSLECTGFKETNLGVYFLSSKWVLRLYALHIID

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YSAVLFPC*AMDHLESFIAECDRRTELAKKRLAETQEEISAEVSAKAEKVHELNEEIGKLLAKAEQLGAEGNVDESQKILMEVEKVRAKKKEAEKTVAEKQEKRNQDRLRRREERERERLSRRSGSRTRDRRRSRSRDRRRRRSRSTSRERRKLSRSRSRDRHRRHRSRSRSHSRGHRRASRDRSAKYKFSRERASREESWESGRSERGPPDWRLESSNGKMASRRSEEKEAG/GDLLNRMIVWKHGLLI
3371	A	345	1383	DLSLECTGFKETNLGVYFLSSKWVLRLYALHIIDYSAVLFPC*AMDHLESFIAECDRRTELAKKRLAETQEEISAEVSAKAEKVHELNEEIGKLLAKAEQLGAEGNVDESQKILMEVEKVRAKKKEAEKTVAEKQEKRNQDRLRRREERERERLSRRSGSRTRDRRRSRSRDRRRRRSRSTSRERRKLSRSRSRDRHRRHRSRSRSHSRGHRRASRDRSAKYKFSRERASREESWESGRSERGPPDWRLESSNGKMASRRSEEKEAG/GDLLNRMIVWKHGLLI
3372	A	239	3348	PMQNCMCSLTLSVLPLGPQPPVPEKRPPEIQHFRMSDDVHSLGKVTSDLAKRRKLTS*GGLSEELGSARRSGEVTLTKGDPGSLEEWETVVGDDFSLYYDSYSVDERVDSDSKSEVEALTEQLSEEEEEEEEEEEEEEEEEEEEEEDEESGNQSDRSGSSGRRKAKKKWRKDSPWVKPSRKRRKREPPRAKEPRGVNGVGSSGPSEYMEVPLGSLELPSEGTLSPNHAGVSNDTSSLETERGFEELPLCSCRMEAPKIDRISERAGHKCMATESVDGELSGCNAAILKRETMRPSSRVALMVLCETHRARMVKHHCCPGCGYFCTAGTFLECHPDFRVAHRFHKACVSQLNGMVFCPHCGEDASEAQEVTIPRGDGVTPPAGTAAPAPPPLSQDVPGRADTSQPSARMRGHGEPRRPPCDPLADTIDSSGPSLTLPNGGCLSAVGLPLGPGREALEKALVIQESERRKKLRFHPRQLYLSVKQGELQKVILMLLDNLDPNFQSDQQSKRTPLHAAAQKGSVEICHVLLQAGANINAVDKQQRTPLMEAVVNNHLEVARYMVQRGGCVYSKEEDGSTCLHHAAKIGNLEMVSLLLSTGQVDVNAQDSGGWTPIIWAAEHKHIEVIRMLLTRGADVTLTDNEENICLHWASFTGSAAIAEVLLNARCDLHAVNYHGDTPLHIAARESYHDCVLLFLSRGANPELRNKEGDTAWDLTPERSDVWFALQLNRKLRLGVGNRAIRTEKIICRDVARGYENVPIPCVNGVDGEPCPEDYKYISENCETSTMNIDRNITHLQHCTCVDDCSSSNCLCGQLSIRCWYDKDGRLLQEFNKIEPPLIFECNQACSCWRNCKNRVVQSGIKVRLQLYRTAKMGWGVRALQTIPQGTFICEYVGELISDAEADVREDDSYLFDLDNKDGEVYCIDARYYGNISRFINHLCDPNIIPVRVFMLHQDLRFPRIAFFSSRDIRTGEELGFDYGDRFWDIKSKYFTCQCGSEKCKHSAEAIALEQSRLARLDPHPELLPELGSLPPVNT
3373	A	587	1584	PDGRLIVSCSEDKTIKIWDTTNKQCVNNFSDSVGFANFVDFNPSGTCIASAGSDQTVKVWDVRVNKLLQHYQVHSGGVNCISFHPSGNYLITASSDGTLKILDLLKGRLIYTLQGHTGPVFTVSFSKGGELFASGGADTQVLLWRTNFDELHCKGLTKRNLKRLHFDSPPHLLDIYPRTPHPHEEKVETVEDFFLHLLRLIQSLR*SICRSLLPLLWISFLLILPQQQKPVVGLCQTRV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KRPVDIS*TLP*CHQNVCQQPRKRKQKT*VTSPV KVK/VSIPLAVTDALEHIMEQLNVLTQTVSILEQR LTLTEDKLKDCLENQQKLFSAVQQKS
3374	A	398	21	WLYPMALSILDIKMSPSWYFHMAIGIINWNTTAG LSGTLYPKVPQKYILFDSVILLLGMLRKIRQVCQ NVYMKGCSPITLFKIVHYWPGAVAHAYNPSTLG GQVG/WQIT*GQEFETSLDYMVKPHLY
3375	A	3	1051	VPTQQILAFPEQTNTKDWTVTPEHVLPESQSLLT FEEVAMYFSQEEWELLDPTQKALYNDVMQENY ETVISLALFVLPKPKVISCLEQGEEPWVQVSPEFK DSAGKSPTGLKLKNDTENHQPVSLSDLEIQASAG VISKKAKVKVPQKTAGKENHFDMHRVGKWHQ DFPVKKRKKLSTWKQELLKLMDRHKKDCAREK PFKCQECGKTFRVSS\DL\IKHQRIHTEEKPYKCQ QCDKRFRWSSDLNKHLTTHQGIKPYKCSWGGKS FSQNTNLHTHQRTHTGEKPFTCHECGKKFSQNS HLIKHRRTHTGEQPYTCSICRRNFSRRSSLLRHQK LHL*REACPVSHFWKTF
3376	A	137	2329	SFESPAPLPSTCFPQERQDPGPCYVSGAMAGLGP GVGDSEGGPRPLFCRKGALRQKVVHEVKSHKFT ARFFKQPTFCSHCTDFIWGIGKQGLQCQVCSFVV HRRCHEFVTFECPGAGKGPQTDDPRNKHKFRLH SYSSPTFCDHCGSLLYGLVHQGMKCSCCEMNVH RRCVRSVPSLCGVDHTERRGRLQLEIRAPTADEI HVTVGEARNLIPMDPNGLSDPYVKLKLIPDPRNL TKQKTRTVKATLNPVWNETFVFNLKPGDVERRL SVEVWDWDRTSRNDFMGAMSFGVSELLKAPVD GWYKLLNQEEGEYYNVPVADADNCSLLQKFEA CNYPLELYERVRMGPSSSPIPSPSPSPTDPKRCFFG ASPGRLHISDFSFLMVLGKGSFGKVMLAERRGSD ELYAIKILKKDVIVQDDDVDCTLVEKRVLALGG RGPGGRPHFLTQLHSTFQTPDRLYFVMEYVTGG DLMYHIQQLGKFKEPHAAFYAAEIAIGLFFLHNQ GIIYRDLKLDNVMLDAEGHIKITDFGMCKENVFP GTTTRTFCGTPDYIAPEIIAYQPYGKSVDWWSFG VLLYEMLAGQPPFDGEDEEELFQAIMEQTVTYP KSLSREAVAICKGFLTKHPGEAPGASGP*WGNLT IRAHGFFPLGFDWERLERL\EIPASFSRPRPCGPQR RGIFDKFFTRAAPA\LTPPARLVLDSIDQADFQGF TYVNPDFVQPDARSPTSTVHVPVM
3377	A	918	738	SSMLWGFSVFRRSWILNCWLSSSQVGISAACKFS TLTHTHTHTHTHTRHAPFCGTCLYY
3378	A	1126	456	FSKLIMKTFIIGISGVTNSGKTTLAKNLQKHLPNC SVISQDDFFKPESEIETDKNGFLQYDVLEALNME KMMSAISCWMESARHSVVSTDQESAEEIPILIIEG FLLFNYKPLDTIWNRSYFLTIPYEECKRRRSTRVY QPPDSPGYFDGHVWPMYLKYRQEMQDITWEVV YLDGTKSEEDLFLQVYEDLIQELAKQKCLQVTA* RRNTTNPS/CK*IRKLQGVI
3379	A	1126	456	FSKLIMKTFIIGISGVTNSGKTTLAKNLQKHLPNC SVISQDDFFKPESEIETDKNGFLQYDVLEALNME KMMSAISCWMESARHSVVSTDQESAEEIPILIIEG FLLFNYKPLDTIWNRSYFLTIPYEECKRRRSTRVY QPPDSPGYFDGHVWPMYLKYRQEMQDITWEVV YLDGTKSEEDLFLQVYEDLIQELAKQKCLQVTA*

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RRNTTNPS/CK*IRKLQGVI
3380	A	1443	794	ARRGELAGGGRASGGRSGGDGGGGGGARAPEG VRAPAAGQPRATKGAPPPPGTPPPSPMSSAIERKS LDPSEEPVDEVLQIPPSLLTCGGCQQNIGDRYFLK AIDQYWHEDCLSCDLCGCRLGEVGRRLYYKLGR KLCRRDYLRLFGQDGLCASCDKRIRAYEMTMRV KDKVYHLECFKCAACQKHFCVGDRYLLINSDIV CEQDIYEWTKINGMI
3381	A	945	474	SLKLRKPPLPTDGVHFVFVESQLDFWGPQEMLT QQGMALQNYDNKLVKCIEELCQKQEELCWQIQ QEEDKKQRLQNEVRQLTEKLACVNEKLARVNE NLARKIASCSKFYQTIAETEATYLKILESF*\TLLS VRKREAGNLTKATAPDQKSSGGRDS
3382	A	1	1458	GIRGKMADRGGVGEAAAVGASPASVPGLNPTLG WRERLRAGLAGTGASLWFVAGLGLLYALRIPLR LCENLAAVTVFLNSLTPKFYVALTGTSSLISGLIFI FEWWYFHKHGTSFIEQVSVSHLQPLMGGTESSIS EPGSPSRNRENETSRQNLSECKVWRNPLNLFRGA EYRRYTWVTGKEPLTYYDMNLSAQDHQTFFTC DTDFLRPSDTVMQKAWRERNPPARIKAAYQALE LN/E*LCHCICSTG*GRSNNYCRC*KVI*TGTQGR RNNL*AVTAVPAPKSSA*SSTEERYQCTGIY*LKI GNVCKKIRKNKRSSKNNERFDE*ISSSYHVEHP* KSL\KSLLELQAYPDVQAVLAKYDDISLPKSAAIC YTAALLKTRTVSEKFSPETASTRGLSAAEINAVD AIHRAVEFNPHVPKYLLEMKSLILPPEHILKRGDS EAIAYAFFHLQHWKRIEGALNLLQCTWEGSKYS FPKVTLISLTIH
3383	A	282	2443	RGKGFKEFFLGVCQTFIPCLCAEGIQLQFFCSGSG SSPLLKDLESMKTGLFFLCLLGTAAAIPTNARLLS DHSKPTAETVAPDNTAIPSLRAEAEENEKETAVS TEDDSHHKAEKSSVLKSKEESHEQSAEQG\KSS\S QELGIEGFKRDSDGSL*VWNL\EYGTNLKGTLDI KEDMSEPQEKKLSENTDFLAPGVSSFTDSNQQES ITKREENQEQPRNYSHHQLNRSSKHSQGLRDQG NQEQDPNISNGEEEEEKEPGEVGTHNDNQERKTE \LPREHANSKQEEDNTQSDDILEESDQPTQVSKM QEDEFDQGNQEQEDNSNAEMEEENASNVNKHIQ ETEWQSQEGKTGLEAISNHKETEEKTVSEALLME PTDDGNTTPRNHGVDDDGDDDGDDGGTDGPRH SA\SDDYFHPKPGLFWEAERA\HSIAYSPSKLREQ REKVHENENIGTTEPGEHQEAKKAENSSNEEETS SEGNMR\VHAVDSCMSFQCKRGHICKADQQGKT SLVSCQDPVT\CPPTKPLDQVCGTDNQTYASSCH LFATKCRLEGTKKGHQLQLDYFG\ASKSIPT\CRD FEVIQ\FPLRMRDW\LKNILMQLYEANSEHAGYL NEK\QRNKVKKIYL\DEKRLLAGDHPIDLLLRDFK KNYHMYVYPVHWQFSELDQHPMDRVLTHSELA PLRASLVPMEHCITRFFEECDPNKDKHITLKEWG HCFGIKEEDIDENLLF
3384	A	3166	928	PSRPHPTHAAMAGPEGFQYRALYPFRRERPEDLE LLPGDVLVVSRAALQALGVAEGGERCPQSVGW MPGLNERTRQRGDFPGTYVEFLGPVALARPGPR PRGPRPLPARPRDGAPEPGLTLPDLPEQFSPPDVA PPLLVKLVEAIERTGLDSESHYRPELPAPRTDWSL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SDVDQWDTAALADGIKSFLLALPAPLVTPEASAE ARRALREAAGPVGPALEPPTLPLHRALTLRFLLQ HLGRVASRAPALGPAVRALGATFGPLLLRAPPPP SSPPPGGAPDGSEPSPDFPALLVEKLLQEHLEEQE VAPPALPPKPPKAK\PASTVPGPNGGSPPSL\QDA EWYWGD\ISREEVNEKLRDTPDGTFLVRDASSKI QGEYTLTLRKGGNNKLIKVFHRDGHYGFSEPLTF CSVVDLINHYRHESLAQYNAKLDTRLLYPVSKY QQDQIVKEDSVEAVGAQLKVYHQQYQDKSREY DQLYEEYTRTSQELQMKRTAIEAFNETIKIFEEQG QTQEKCSKEYLERFRREGN/QTKEMQRILLNSER LKSRIA\EIHESRT\KL\EQQLLVPRASDNKRD/IDK PH*TSLKPDLMQLRKIRDQYLVWLTQKGARQKK INEWLGIKNETEDQYALMEDEDDLPHHEERTWY VGKINRTQAEEMLSGKRDGTFLIRESSQRGCYAC SVVVDGDTKHCVIYRTATGFGFAEPYNLYGSLK ELVLHYQHASLVQHNDALTVTLAHPVRAPGPGP PPAAR
3385	A	43	2372	TRDVNSWKELCFNHYNKETTNCYRTTRKWTNY KIIFLGPFRELRSQGNQVILNLGKERCQLRETGLK LYLPGMDSARHHISHSTSAGPIPSQKEEEMTESQ GTVTFKDVAIDFTQEEWKRLDPAQRKLYRNVML *NYNNLITVGYPFTKPDVIFKLEQEEKPWVMEEE VLRRHWQGEIWGVDEHQKNQDRLLRQVEVKFQ KTLTEEKGNECQKKFANVFPLNSDFFPSRHNLYE YDLFGKCLEHNFDCHNNVKCLMRKEHCEYNEP VKSYGNSSSHFVITPFKCNHCGKGFNQTLDLIRH LRIHTGEKPYECSNCRKAFSHKEKLIKHYKIHSRE QSYKCNECGKAFIKMSNLIRHQRIHTGEKPYACK ECEKSFSQKSNLIDHEKIHTGEKPYECNECGKAFS QKQSLIAHQKVHTGEKPYACNECGKAFPRIASLA LHMRSHTGEKPYKCDKCGKAFSQFSMLIIHVRIH TGEKPYECNECGKAFSQSSALTVHMRSHTGEKP YECKECRKAFSHKKNFITHQKIHTREKPYECNEC GKAFIQMSNLVRHQRIHTGEKPYICKECGKAFSQ KSNLIAHEKIHSGEKPYECNECGKAFSQKQNFIT HQKVHTGEKPYDCNECGKAFSQIASLTLHLRSHT GEKPYECDKCGKAFSQCSLLNLHMRSHTGEKPY VCNECGKAFSQRTFLIVHMRGHTGEKPYECNEC GKAFSQSSSLTIHIRGHTGEKPYECKECRKAFSHK KNFITHQKIHTRE/KPFKCNHCGKGFNQTLDLIRH LRIHTGEKPYECSNCRKAFSHKEKLIKHYKIHSRE QSYKCNECGKAFIKMSNLIRHQRIHTGEKPYACK ECEKSFSQKSNLIDHEKIHTGEKPYECNECGKAFS QKQSLIAHQKVHTGEKPYACNECGKAFPRIASLA LHMRSHTGEKPYKCDKCGKAFSQFSMLIIHVRIH TGEKPYECNECGKAFSQSSALTVHMRSHTGEKP YECKECRKAFSHKKNFITHQKIHTREKPYECNEC GKAFIQMSNLVRHQRIHTGEKPYICKECGKAFSQ KSNLIAHEKIHSGEKPYECNECGKAFSQKQNFIT HQKVHTGEKPYDCNECGKAFSQIASLTLHLRSHT GEKPYECDKCGKAFSQCSLLNLHMRSHTGEKPY VCNECGKAFSQRTFLIVHMRGHTGEKPYECNEC GKAFSQSSSLTIHIRGHTGEKPYECKECRKAFSHK KNFITHQKIHTRENPLSVIIVEKASIRLWTSSDI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3386	A	201	1032	WDDYPQGALRRREAAEGLHFLGPPGRVRGQLR GITGPAWYCHSPSHSLLSAFCHLPTPSRCPAMAR PPVPGSVVVPNWHES/RRGQGVPGLHSAQEPPAG VWAA*AASAAAA\LSIDTASYKIFVSGKSGVGKT ALVAKLAGLEVPVVHHETTGIQTTVVFWPAKLQ ASSRVVMFRFEFWDCGESALKKFDHMLLACME NTDAFLFLFSFTDRASFEDLPGQLARIAGEAPGV VRMVIGSKFDQYMHTDVPERDLTAFRQAWELPL LRVKSVPGRRLG
3387	A	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSPGQP EAGPEGAQERPSQAAPAVEAEGPGSSQAPRKPEG AQARTAQSGALRDVSEELSRQLEDILSTYCVDNN QGGPGEDGAQGEPAEPEDAEKSRTYVARNGEPE PTPVVNGEKEPSKGDPNTEEIRQSDEVGDRDHRR PQEKKKAKGLGKEITLLMQTLNTLSTPEEKLAAL CKKYAELLEEHRNSQKQMKLLQKKQSQLVQEK DHLRGEHSKAVLARSKLESLCRELQRHNRSLKE EGVQRAREEEEKRKEVTSHFQVTLNDIQLQMEQ HNERNSKLRQENMELAERLKKLIEQYELREEHID KVFKHKDLQQQLVDAKLQQAQEMLKEAEERHQ REKDFLLKEAVESQRMCELMKQQETHLKQQLA LYTEKFEEFQNTLSKSSEVFTTFKQEMEKMTKKI KKLEKETTMYRSRWESSNKALLEMAEEKTVRD KELEGLQVKIQRLEKLCRALQT/GAQ*PVRGQRW GSHRTSAVRIFS
3388	A	98	3197	ARPEVPAPPAWLSRRGAAKMGDKKDDKDSPKK NKGKERRDLDDLKKEVAMTEHKMSVEEVCRKY NTDCVQGLTHSKAQEILARDGPNALTPPPTTPEW VKFCRQLFGGFSILLWIGAILCFLAYGIQAGTEDD PSGDNLYLGIVLAAVVIITGCFSYYQEAKSSKIME SFKNMVPQQALVIREGEKMQVNAEEVVVGDLV EIKGGDRVPADLRIISAHGCKVDNSSLTGESEPQT RSPDCTHE\NPLKTRNITFFSNNFVEGTARGVVVA TGDRTVMGRIATLASGLEVGKTPIAIEIEHFIQLIT GVAVFLGVSFFILSLILGYTWLEAVIFLIGIIVANV PEGLLATVTVCLTLTAKRMARKNCLVKNLEAVE TLGSTSTICSDKTGTLTQNRMTVAHMWFDNQIH EADTTEDQSGTSFDKSSHTWVALF*H/LLGFCNR PVFKGGQDNIPVLKRDVAGDASESALLKCIELSS GSVKLMRERNKKVAEIPFNSTNKYQLSIHETEDP NDNRYLLVMKGAPERILDRCSTILLQGKEQPLDE EMKEAFQNAYLELGGLGERVLGFCHYYLPEEQF PKGFAFDCDDVNFTTDNLCFVGLMSMIGPPRAA VPDAVGKCRSAGIKVIMVTGDHPITAKAIAKGV GIIFEGNETVEDIAARLNIPVSQVNPRDAKACVIH GTDLKDFTSEQIDEILQNHTEIVFARTSPQQKLIIV EGCQRQGAIVAVTGDGVNDSPALKKADIGVAM GIAGSDVSKQAADMILLDDNFASIVTGVEEGRLI FDNLKKSIAYTLTSNIPEITPFLLFIMANIPLPLGTI TILCIDLGTDMVPAISLAYEAAESDIMKRQPRNPR TDKLVNERLISMAYGQIGMIQALGGFFSYFVILA ENGFLPGNLVGIRLNWDDRTVNDLEDSYGQQW TYEQRKVVEFTCHTAFFVSIVVVQWADLIICKTR RNSVFQQGMKNKILIFGLFEETALAAFLSYCPGM DVALRMYPLKPSWWFCAFPYSFLIFVYDEIRKLI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LRRNPGGWVEKETYY
3389	A	45	5250	VERLLGCRNSKRTWRMLISKNMPWRRLQGISFG MYSAEELKKLSVKSITNPRYLDSLGNPSANGLYD LALGPADSKEVCSTCVQDFSNCSGHLGHIELPLT VYNPLLFDKLYLLLRGSCLNCHMLTCPRAVIHLL LCQLRVLEVGALQAVYELERILNRFLEENPDPSA SEIREELEQYTTEIVQNNLLGSQGAHVKNVCESK SKLIALFWKAHMNAKRCPHCKTGRSVVRKEHNS KLTITFPAMVHRTAGQKDSEPLGIEEAQIGKRGY LTPTSAREHLSALWKNEGFFLNYLFSGMDDDGM ESRFNPSVFFLDFLVVPPSRYRPVSRLGDQMFTN GQTVNLQAVMKDVVLIRKLLALMAQEQKLPEE VATPTTDEEKDSLIAIDRSFLSTLPGQSLIDKLYNI WIRLQSHVNIVFDSEMDKLMMDKYPGIRQILEK KEGLFRKHMMGKRVDYAARSVICPDMYINTNEI GIPMVFATKLTYPQPVTPWNVQELRQAVINGPN VHPGASMVINEDGSRTALSAVDMTQREAVAKQ LLTPATGAPKPQGTKIVCRHVKNGDILLLNRQPT LHRPSIQAHRARILPEEKVLRLHYANCKAYNADF DGDEMNAHFPQSELGRAEAYVLACTDQQYLVP KDGQPLAGLIQDHMVSGASMTTRGCFFTREHYM ELVYRGLTDKVGRVKLLSPSILKPFPLWTGKQVV STLLINIIPEDHIPLNLSGKAKITGKAWVKETPRSV PGFNPDSMCESQVIIREGELLCGVLDKAHYGSSA YGLVHCCYEIYGGETSGKVLTCLARLFTAYLQL YRGFTLGVEDILVKPKADVKRQRIIEESTHCGPQ AVRAALNLPEAASYDEVRGKWQDAHLGKDQRD FNMIDLKFKEEVNHYSNEINKACMPFGLHRQFPE NTLQLMVQSGAKGSTVNTMQISCLLGQIELEGRS TPLMASGKSLPCFEPYEFTPRAGGFVTGRFLTGIK PPEFFFHCMAGREGLVDTAVKTSRSGYLQRCIIK HLEGLVVQYDLTVRDSDGSVVQFLYGEDGLDIP KTQFLQPKQFPFLASNYEVIMKSQHLHEVLSRAD PKKALHHFRAIKKWQSKHPNTLLRRGAFLSYSQ KIQEAVKALKLESENRNGR/RPWDS/G/RMLRMW YELDEESRRKYQKKAAACPDPSLSVWRPDIYFAS VSETFETKVDDYSQEWAAQTEKSYEKSELSLDR LRTLLQL\KWQRSLCEPGEAVGLLAAQSIGEPST QMTLNTFHFAGRGEMNVTLGIPRLREILMVASA NIKTPMMSVPVLNTKKALKRVKSLKKQLTRVCL GEVLQKIDVQESFCMEEKQNKFQVYQLRFQFLP HAYYQQEKCLRPEDILRFMETRFFKLLMESIKKK NNKASAFRNVNTRRATQRDLDNAGELGRSRGE QEGDEEEEGHIVDAEAEEGDADASDAKRKEKQE EEVDYESEEEEEREGEENDDEDMQEERNPHREG ARKTQEQDEEVGL/GH*GGPVPSRPPDAAPETHP QPGAPGA\EAMERRVQAVREIHPFIDDYQYDTEE SLWCQVTVKLPLMKINFDMSSLVVSLAHGAVIY ATKGITRCLLNETTNNKNEKELVLNTEGINLPELF KYAEVLDLRRLYSNDIHAIANTYGIEAALRVIEK EIKDVFAVYGIAVDPRHLSLVADYMCFEGVYKP LNRFGIRSNSSPLQQMTFETSFQFLKQATMLGSH DELRSPSACLVVGKVVRGGTGLFELKQPLR
3390	A	2	2080	ILPPLEGPPAQASPSSTMLGEGSQPDWPGGSRYD LDEIDAYWLELINSELKEMERPELDELTLERVLE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ELETLCHQNMARAIETQEGLGIEYDEDVVCDVC RSPEGEDGNEMVFCDKCNVCVHQACYGILKVPT GSWLCRTCALGVQPKCLLCPKRGGALKPTRSGT KWVHVSCALWIPEVSIGCPEKMEPITKISHIPASR WALSCSLCKECTGTCIQCSMPSC\VTAFHVTCAF DHGLEMRTILADNDEVKFKSFCQEHSDGGPRNE PTSEPTEPSQAGEDLEKVTLRKQRLQQLEEDFYE LVEPAEVAERLDLAEALVDFIYQYWKLKRKANA NQPLLTPKTDEVDNLAQQEQDVLYRRLKLFTHL RQDLERVRNLCYMVTRRERTKHAICKLQEQIFH LQMKLIEQDLCRAGLSTSFPIDGTFFNSWLAQSV QITAENMAMSEWPLNNGHREDPAPGLLSEELLQ DEETLLSFMRDPSLRPGDPARKARGRTRLPAKK KPPPPPPQDGPGSRTTPDKAPKKTWGQDAGSGK GGQGPPTRKPPRRTSSHLPSSPAAGDCPILATPES PPPLAPETPDEAASVAADSDVQVP\GPAASPKPLG RLRPPPREPR*T\RRLPGC/ARPDAGDGDHLSAVA ERPKV\SLHFDTETDG\YFS\DGEMSNS\DV\EAED GGVQRGPREAGAKE\VVRMGVLAS
3391	A	1555	327	NSFLHFLHLKVRTMFLFPSFPVLLLSVVTASCSKT KACADTQKTCSMITCGIPVTNGTPGRDGRDRPK GEKGEPGLGQVSVAS*ISTSGRCSSKSVLEPATRG LKHRLGEAPLSSGPMLHSEQPL*NAIASKTKLFV DSLGSHISTQELGVCGCPFRGVSCLVGELALVQA LH*VAGESFFFGSDHWLIGCAGGEQEWSIELLGK KKRVTATGSSSLCLATGQGLRGLQGPPGKMGPP GNTGTSGIPGPRGQKGDRGDNSVAEAKLANLER KL*SLRSELDHTKKL*PFSLGK\MSGKKLFVTNGE RMPFSKVKALCAGLQATVAAPKNAEENKAIQDV AKDTAFLGITDEATEGQFMYLTGGRLTYSNWKK DEPNDHGSGEDCVILLNNGLWNGISCTSSFIAICE FPA
3392	A	218	1773	GGSRRNQRRSIPVLGYFLKQKKMTKAQESLTLE DVAVDFTWEEWQFLSPAQKDLYRDVMLENYSN LVSVGYQAGKPDALTKLEQGEPLWTLEDEIHSP AHPEIEKADDHLQQPLQNQKILKRTGQRYEHGR TLKSYLGLTNQSRRYNRKEPAEFNGDGAFLHDN HEQMPTEIEFPESRKPISTKSQFLKHQQTHNIEKA HECTDCGKAFLKKSQLTEHKRIHTGKKPHVCSL CGKAFYKKYRLTEHERAHRGEKPHGCSLCGKAF YKKRYRLTEHERAHKGEKPYGCSECGKAFPRKSE LTEHQRIHTGIKPHQCSECGRAFSRKSLLVVHQR THTGEKPHTCSECGKGFIQKGNLNIHQRTHTGEK PYGCIDCGKAFSQKSCLVAHQRYHTGKTPFVCPE CGQPCSQKSGLIRHQKIHSGEKPYKCSDCGKAFL TKTMLIVHHRTHTGERPYGCDECEKAYFYMSCL VKHKRIHSREKRGD/CSEGGKSFHSKSQLKS**TC AGEKPC*YGNCGNGGRAV
3393	A	46	1464	ARSLSGAPSGSSRQDGTSLLRTGAGYSSSQSIETL SLPPGPSHLVGDKSQGGRSCQGQITSAASGKTSK SEPNHVIFKKISRDKSVT\IYLGNRDY\IDHV\SQV QPVDGVVLVDPDLVKGKKVYVTLTCAFRYGQE DIDVIGLTFRRDLYFSRVQVYPPVGAASTPTKLQ ESLLKKLGSNTYPFLLTFPDYLPCSVMLQPAPQD SGKSCGVDFEVKAFATDSTDAEEDKIPKKSSVRL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LIRKVQHAPLEMGPQPRAEAAWQFFMF\DKPLH LAVSLNKRDLFPMGSPIPVPVSVP\NNTEKPVKKI KA\SVEQVANVVLYS\SDY\YVKPVAMEEAQEKV PPNSTWTKA\LTLL\PWLVNNRERRGIALDGKIKH EDTNLASSTIIKEGIDRKRSWEILVSYPDQR*SSTV SGFLGRASPSQ*SRPT*RSQFRL\MHPQP\EDPA\K ESYQDANLVF\EEFARP*ILKDAGEA*\EGKRDQE
3394	A	211	1591	RPPTMAADQRPKADTLALRQRLISSSCRLFFPEDP VKIVRAQGQYMYDEQGAEYIDCISNVAHVGHCH PLVVQAAHEQNQVLNTNSRYLHDNIVDYAQRLS ETLPEQLCVFYFLNSGSEANDLALRLARHYTGH QDVVVLDHAYHGHLSSLIDISPYKFRNLDGQKE WVHVAPLPDTYRGPYREDHP\THVEDGLEKAFS* KRVVQGRNRQICRRQIAAFFAESLPSVGGQIIPPA GYFSQVAEHIRKAGGVFVADEIQVGFGRVGKHF WAFQLQGKDFVPDIVTMGKSIGNGHPVACVAAT QPVARAFEATGVEYFNTFGGSPVSCAVGLAVLN VLEKEQLQDHATSVGSFLMQLLGQQKIKHPIVG DVRGVGLFIGVDLIKDEATRTPATEEAAYLVSRL KENYVLLSTDGPGRNILKFKPPMCFSLDNARQV VAKLDAILTDMEEKVRSCETLRLQP
3395	A	1	1424	FRDGFSLRCGCNAELPGRGGDDAADRAIQRFLR TGAAVRYKVMKNWGVIGGIAAALAAGIYVIWG PITERKKRRKGLVPGLVNLGNTCFMNSLLQGLSA CPAFIRWLEEFTSQYSRDQKEPPSHQYLSLTLLHL LKALSCQEVTDDEVLHASCLLDVLRMYRWQISS FEEQDAHELFHVITSSLEDERDRQPRVTHLFDVH SLE\HSQK*LPKQITCRTRGSPHPTSNHWKSQHPF HGRLTSNMVCKHCEHQSPVRFDTFDSLSLSIPAA TWGHPLTLDHCLHHFISSESVRDVVCDNCTKIEA KGTLNGEKVEHQRTTFVKQLKLGKLPQCLCIHL QRLSWSSHGTPLKRHEHVQFNEFLMMDIYKYHL LGHKPSQHNPKLNKNPGPTLELQDGPGAPTPGL NQPGAPKTQIFMNGACSPSLLPTLSAPMPFPLPV VPDYSSSTYLFRLMGSCRPPWETWHSGTLCSFTD GPHL
3396	A	109	107	TQEAGLIFFSPPFSLSLSLSLPLSLFLLSHPHSRTPP NRTPRRTRIPQRPAVMYSPLCLTQDEFHPFIEALL PHVRAFAYTWFNLQARKRKYFKKHEKRMSKEE ERAVKDELLSEKPEVKQKWASRLLAKLRKDIRP EYREDFVLTVTGKKPPCCVLSNPDQKGKMRRID CLRQADKVWRLDLVMVILFKGIPLESTDGERLV KSPQCSNPGLCVQPHHIGVSVKELDLYLAYFVH AADSSQSESPSQAK*R*H*GPARKWDIWGFQ\DS FVT\SGVF\SVT*A*LRVSQTPI\AAG\TGPNFSLSD LESSSYYSMSPGAMRRSLPSTSSTSSTKRLKSVED EMDSPGEEPFYTGQGRSPGSGSQSSGWHEVEPG MPSPTTLKKSEKSGFSSPSPSQTSSLG\TAFTQHHR PVITGTQSKFHIATPSIL\HFPRHSPFFQQPGPYFSH PAIRYHPQETLKEFVQLVCPDAGQQAGQPNGSS QGKVHNPFLPTPMLPPPPPPPMARPVPLPVPDTK PPTTSTEGGAASPTSPTTRS/PGRTRPQQPFL/SYG PP*PSNALIGGGGGGAGERAGERADLEM
3397	A	1	2002	TGTLTEDGLDVMGVVPLKGQAFLPLVPEPRRLP VGPLLRALATCHALSRLQDTPVGDPMDLKMVES

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TGWVLEEEPAADSAFGTQVLAVMRPPLWEPQLQAMEEPPVPVSVLHRFPFSSALQRMSVVVAWPGATQPEAYVKGSPELVAGLCNPETVPTDFAQMLQSYTAAGYRVVALASKPLPSVPSLEAAQQLTRDTVEGDLSLLGLLVMRNLLKPQTTPVIQALRRTRIRAVMVTGDNLQTAVTVARGCGMVAPQEHLIIVHATHPERGQPASLEFLPMESPTAVNGVKDPDQAASYTVEPDPRSRHLALSGPTFGIIVKHFPKLLPKVLVQGTVFARMAPEQKTELVCELQKLQYCVGMCGDGANDCGALKAADVGISLSQAEASVVSPFTSSMASIECVPMVIREGRCSLDTSFSVFKYMALYSLTQFISVLILYTINTNLGDLQFLAIDLVITTTVAVLMSRTGPALVLGRVRPPGALLSVPVLSSLLLQMVLVTGVQLGGYFLTLAQPWFVPLNRTVAAPDNLPNYENTVVFSLSSFQYLILAAAVSKGAPFR\RPLTNNVPFLLASAL*SSVLVVLVLSPGLLHGPLALRNITDTGFKLLLVGLVTLNFVGGLHAGERARPVPPRLPAPPPAQAG\SKKRFKQLERELAEQPWPPLPAGPLR
3398	A	758	1368	FPFRMLTGYLYLMWRRKAFWSGTQRHPLPGGLKRRRRPGRGPWPAPGGQGVGPSAL*KAGSPPANRPGQGE/PGLISPKPVTEVLPDVQGAPVPVPPLPTPPSLPHLQNQPP/TVQHYLLSFSWKPSQGPE*RA*PSPLPPAAMRPDG*PGPASQGPDQPG\PCPPASLPTSPPGKGFQKTETRKHPPPRQQHKPKCTANRPLASFL
3399	A	906	1091	HHHHHHHHHHHHHLVAFGKVQ*LQNSPSSSSSSSGCFWQARFSSYRTLHHHHHHHHHHHHH
3400	A	1838	325	PFLSVHRSPHGPSKLCDDPQASLVPEPVPGGCQEPEEMSWPPSGEIASPPELPSSPPPGLPEVAPDATSTGLPDTPAAPETSTNYPVECTEGSAGPQSLPLPILEPVKNPCSVKDQTPLQLSVEDTTSPNTKPCPPTPTTPETSPPPPPPPPSSTPCSAHLTPSSLFPSSLESSSEQKFYNFVILHARADEHIALRVSGRSWEALGVPDGATFCEDFQVPGRGELSCLQDAIDHSAFIILLLT\SN\FDCR\LSLHQVNQAMMSNLT\RQGSQDCVIP\FLP\LESSPARLSSDTASLLSGLVRLDEHSQIFARKVANTFKPHRLQARKAMWRKEQDTRALREQSQHLDGERMQAAALNAAYSAYLQSYLSYQAQMEQLQVAFGSHMSFGTGAPYGARMPFGGQVPLGAPPPFPTWPGCPQPPPLHAWQAGTPPPPSPQPAAFPQSLPFPAVPKPFPTASTAPPSEPKGWQP\LIIHHAQMVTSWG*NKH\MWNQRGSQAPEDKTQEAE
3401	A	153	1389	EWGWLGAAQPPEEEAEAEDQESPSSLCREALAEIKKEISPLFIGMEKCSVGGLELTEQTPALLGNMAMATSLMDIGDSFGHPACPLVSRSRNSPVEDDDDDDDVVFIESIQPPSISAPAIADQRNFIFASSKNEKPQGNYSVIPPSSRDLASQKGNISETIVIDDEEDIETNGGAEKKSSCFIEWGLPGTKNKTNDLDFSTSSLSRSKVNAGMGNSGITTELTLKYIITNVTTLETGISSVNAGQDVNIIITYKTSL*NTNLGDVAKGLQSSNFGVNIQTYTPSLTPQTKTGV\NLLTLVE*MWQETYFRMENLQLII/CPEDASTKKANVILPVESSKSFQEFYSTSCLSPCENNWNLKKGVFNKSRCTICSKLAEVWIFIPKLLFRLTVIILTFKCYYVLFHLHNARVLDV
3402	A	153	1389	EWGWLGAAQPPEEEAEAEDQESPSSLCREALAEI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KKEISPLFIGMEKCSVGGLELTEQTPALLGNMAM ATSLMDIGDSFGHPACPLVSRSRNSPVEDDDDDD DVVFIESIQPPSISAPAIADQRNFIFASSKNEKPQG NYSVIPPSSRDLASQKGNISETIVIDDEEDIETNGG AEKKSSCFIEWGLPGTKNKTNDLDFSTSSLSRSK VNAGMGNSGITTELTLKYIITNVTTLETGISSVNA GQDVNIIITYKTSL*NTNLGDVAKGLQSSNFGVNI QTYTPSLTPQTKTGV\NLLTLVE*MWQETYFRME NLQLII/CPEDASTKKANVILPVESSKSFQEFYSTS CLSPCENNWNLKKGVFNKSRCTICSKLAEVWIFI PKLLFRLTVIILTFKCYYVLFHLHNARVLDV
3403	A	609	2765	SRHCTPAERQNETHRAPDFAMSAVLGHQPPFFPA LTLPPNGAAALSLPGALAKPIMDQLVGAAETGIP FSSLGPQAHLRPLKTMEPEEEVEDDPKVHLEAKE LWDQFHKRGTEMVITKSGRRMFPPFKVRCSGLD KKAKYILLMDIIAADDCRYKFHNSRWMVAGKA DPEMPKRMYIHPDSPATGEQWMSKVVTFHKLKL TNNISDKHGFTILNSMHKYQPRFHIVRANDILKLP YSTFRTYLFPETEFIAVTAYQNDKITQLKIDNNPF AKGFRDTGNGRREKRKQLTLQSMRVFDERHKK ENGTSDESSSEQAAFNCFA\QASSPAA\PL*RTSNL KDF\SPSRG*RATPEAEEQRGSTAPRPATRAKISP HPRRRSPAVTRAAPAVKAHLFAAERPRDSGRLD KASPDSRHSPATISSSTRGLGAEERRSPVREG\QA PAKVEEARALPGKEAFAPLTVQTDAAAAHLAQG PLPGLGFAPGLAGQQFFNGHPLFLHPSQFAMGG AFSSMAAAGMGPLLATVSGASTGVSGLDSTAM ASAAAAQGLSGASAATLPFHLQQHVLASQGLA MSPFGSLFPYPYTYMAAAAAA/SSAAASASVHRT P\FNLNTMRPRLRYSPYSIPVPVPDGSSLLTTALPS MAAAAGPLDGKAAALAASPAS\VAVDSGSELNS RSS\TLSSSSMSLSPKLCAEKEAATSELQSIQRLVS GLEAKPDRSRSASP
3404	A	1082	1308	LKKFLEVPQSYSLLLSSPFLQ\WRA*RPQNAIG*Q FIIKTLVFFGIMRSAGDVLSTQVSCALRIMRTAGC SHSSP
3405	A	1553	559	PRPPTQRLSRFAPPCRTAEFPFRRRAVVTRPAPPR ACTVVGRSSPVTGLAVGAAVAMLTVAARSRPFA PVLSATSRGVAGALT\P*MQATVPATPEQPVLDL KRPFLSRESLSGQAVRRPLVASVGLNVPASVCYS HTDIKVPDFSEYRRLEVLDSTKSSRESSEARKGFS YLVTGVTTVGVAYAAKNAVTQFVSSMSASADV LALAKIEIKLSDIPEGKNMAFKWRGKPLFVRHRT QKEIEQEAAVELSQLRDPQHDLDRVKKPEWVILI GVCTHLGCVPIANAGDFGGYYCPCHGSHYDASG RIRLGPAPLNLEVPTYEFTSDDMVIVG
3406	A	83	2671	CLYPDFCRSVTCAMPCFTHRSCREDPGTSESREM DPVAFKDVAVNFTQEEWALLDISQKNLYREVML ETFWNLTSIGKKWKDQNIEYEYQNPRRNFRSVT EEKVNEIKEDSHCGETFTPVPDDRLNFQKKKASP EVKSCDSFVCEVGLGNSSSNMNIRGDTGHKACE CQEYGPKPWKSQQPKKAFRYHPSLRTQERDHTG KKPYACKECGKNIIYHSSIQRHMVVHSGDGPYK CKFCGKAFHWLSLYLIHERTHTGEKPYECKQCG KSFSYSATHRIHERTHIGEKPYECQECGKAFHSPR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SCHRHERSHMGEKAYQCKECGKAFMCPRYVRR HERTHSRKKLYECKQCGKALSSLTSFQTHIRMHS GERPYECKTCGKGFYSAKSFQRHEKTHSGEKPY KCKQCGKAFTRSGSFRYHERTHTGEKPYECKQC GKAFRSAPNLQSHGRTHTGEKPYECKECGKAFIF VNNLQSHERTQTHIRIHSGERRYKCKICGKGFYC PKSFQRHEKTHTGEKLYEC/TATFSSSFSSSSSF*Y HERTHTGEKPYKCEQCGKAFRAVSIL*MHGRTH PEEKPYECEQ*RKAFRSAPHL*IRGRTHNGEKPY ACKKCGKPFGSAQNLRIHERTQTHIMHSVERPYK CKICGRGFYSAKSFQTHEKSYTGEKPYECKQCG KAFVSFTSFRYHERTHTGENPYECKQFGKAFRSV KNLRFHKRTHTGEKPCEYMKRLTLEGNTMNAS NVAKLSLLPVLFNIMKEFTLGRNPISVSNVRKPLF LPLLFNIMKGLTWERNPMSVCHVGKPSFLLVPFN IMKGLTLERSPMNISNVGKPSDQPRTFKCMEGLT LEKNPMNVSSMGKRSDLTRFFEYR
3407	A	1426	3	PAAPSGASPGRVCGVETARPLGVQRRQSADEGP PGVAGLRHEPPTVWLGSVAHRGTWVCAHRWFG PAVTRAAQAATMVKLLVAKILCMVGVFFFMLL GSLLPVKIIETDFEKAHRSKKILSLCNTFGGGVFL ATC\LTALLARC*GKSSRRSWSLGHISTDYPL\AE TILLLGFFMTVFLEQLILTFAQENAVLHRPGDLQR RIGRGQRLGV*EPLHGGRAGPRAVRGAPRPRPQP ERAGPLA\PSPVRLLSLAFALSAHSVFEGLALGLQ EEGEKVVSLFVGVAVHETLVPVALGISMAGSAM PLRDAAKLAVTVSPMIPLGIGLGLGIEKAQGVPG SVASVLLQGPGGRHLSLFITFPGKSWPRSWRKKS DRLLKVLF\LVVGYTVLAGMGLPQVVSGLAIVPA AGSPPGAPGRTQAASPGRASPKSEHCGPGPPPVH KGPPGTRLCPRSYTLSLRALLLFKILLSLKSLYQK KK
3408	A	106	4514	EARDRLAQSRAKEKELNSVASELSARQEESEHSH KHLIELRREFKKNVPEEIREMVAPVLKSFQAEVV ALSKRSQEAEAAFLSVYKQLIEAPALWELKLKSR PALGDSRVQQGQHDPKTDNQNTQQKAGFKEGW LAEASEREAFGPGFKDPVPVFEAARSLDDRLQPP SFDPSGQPRRDLHTSWKRNPELLSPKALKATQAE LLELRRKYDEEAASKADEVGLIMTNLEKANQRA EAAQREVESLREQLASVNSSIRLACCSPQGPSGD KVNFTLCSGPRLEAALASKDREILRLLKDVQHLQ SSLQELEEASANQIADLERQLTAKSEAIEKLEEKL QAQSDYEEIKTELSILKAMKLASSTCSLPQGMAK PEDSLLIAKEAFFPTQKFLLEKPSLLASPEEDPSED DSIKDSLGTEQSYPSPQQLPPPPGPEDPLSPSPGQP LLGPSLGPDGTRTFSLSPFPSLASGERLMMPPAAF KGEAGGLLVFPPAFYGAKPPTAPATPAPGPEPLG GPEPADGGGGGAAGPGAEEEQLDTAEIAFQVKE QLLKHNIGQRVFGHYVLGLSQGSVSEILARPKP\ WRKLHG**GKEPFIKMKQFLSDEQNVLALRTIQV RQRGSITPRIRTPETGSDDAIKSILEQAKKEIESQK GGEPKTSVAPLSIANGTTPASTSEDAIKSILEQAR REMQAQQQALLEMEVAPRGRSVPPSPPERPSLAT ASQNGAPALVKQEEGSGGPAQAPLPVLSPAAFV QSIIRKVKSEIGDAGYFDHHWASDRGLLSRPYAS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VSPSLSSSSSSGYSGQPNGRAWPRGDEAPVPPED EAAAGAEDEPPRTGELKAEGATAEAGARLPYYP AYVPRTLKPTVPPLTPEQYELYMYREVDTLELTR QVKEKLAKNGICQRIFGEKVLGLSQGSVSDMLSR PKPWSKLTQKGREPFIRMQLWLSDQLGQAVGQQ PGASQASPTEPRSSPSPPPSPTEPEKSSQEPLSLSLE SSKENQQPEGRSSSSLSGKMYSGSQAPGGIQEIV AMSPELDTYSITKRVKEVLTDNNLGQRLFGESIL GLTQGSVSDLLSRPKPWHKLSLKGREPFVRMQL WLNDPHNVEKLRDMKKLEKKAYLKRRYGLIST GSDSESPATRSECPSPCLQPQDLSLLQIKKPRVVL APEEKEALRKAYQLEPYPSQQTIELLSFQLNLKT NTVINWFHNYRSRMRREMLVEGTQDEPDLDPSG GPGILPPGHSHPDPTPQSPDSETEDQKPTVKELEL QEGPEENSTPLTTQDKAQVRIKQEQMEEDAEEE AGSQPQDSGELDKGQGPPKEEHPDPPGNDGLPK VAPGPLLPGGSTPDCPSLHPQQESEAGERLHPDP LSFKSASESSRCSLEVSLNSPSAASSPGLMMSVSP VPSSSAPISPSPPGAPPAKVPSASPTADMAGALHP SAKVNPNLQRRHEKMANLNNIIYRLERAANREE ALEWEF
3409	A	162	1710	GPLSPGPYQCRPSLPAQLYPQSLMAAATLRTPTQ GTVTFEDVAVHFSWEEWGLLDEAQRCLYRDVM LENLALLTSLDVHHQKQHLGEKHFISNVGRALF VKTCTFHVSGEPSTCREVGKDFLAKLGFLHQQA AHTGEQSNSKSDGGAISHRGKTHYNWGEHTKAF SGKHTLVQQQRTLTTERCYICSECGKSFSKSYSL NDHWRLHTGEKPYECRECGKSFRQSSSLIQHRR GHTAVRPHECDECGKLFSNKSNLIKHRRVHTGE RPYECSECGKSFNQRSALLQHRGVHTGEKPYEC TECGKSFSHNSSLIKHQRIHSG*\RPYECTECGKSF SQNSSLIEHHRVHTGERPYKCSECGKSFRQRSAL LQHRGVPTGERPYECSECGKFFPYSSSLGKHQRV HTGSRPYECSECGKSFTQNSGLIKHRRVHTGEKP YECTE*KKSFSHNSSLIKHQRIHSR*KPYE\CKCG N\R*HPGESP*VHSECQ/KSFS*RPYLIECHTVHKG KTLLICRDVQLI
3410	A	167	789	LCMKGISGGVRVAALAARAEREELPVPAMEPQP TAWGSPHPEAVLQLEVAPESSGPCTDTAKDQQS DKLPDLMPPA\EPLGSALELRASLEIDVAE\RGCE HGPSQQLPRCP*SWAWSEPWCQRPGCAV*APLP Y*REASFIYQSHSPAASGPFHSAGAGAVYLQAGG V/GEQEKEAVRKGSGSSSCSQRGP\PPPGMEVCPL LGFWAICP
3411	A	1040	887	ASLSKPAGISTMPWALILLFLLTHSAVSVVQAGL TQPPSVSKDLR\QTATLTCTGNSNNVGHQGVIWL QQHQGHPPKLLSYRNNNRPSGISERLSAYKSGNA ASLTIYGLQTEHEAD**CRPRRKLIPKTARLFFFFL IDNEEYLLRVY
3412	A	164	83	RRGIPGSASLSLTMCVRSCFQSPRLQWVWRTAFL KHTQRRHQGSHRWTHLGGSTYRAVIFDMGGVLI PSPGRVAAEWEVQNRIPSGTILKALMEGGENGP WMRFMRAEITAEGFLREFGRLCSEMLKTSVPVD SFFSLLTSERVAKQFPVMTEAITQIRAKGLQTAVL SNNFYLPNQKSFLPLDRKQFDVIVESCMEGICKP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DPRIYKLCLEQLGLQPSESIFLDDLGTNLKEAARL GIHTIKVNDPETAVKELEALLGFTLRVGVPNTRP VKKTMEIPKDSLQKYLKDLLGIQTTGPLELLQFD HGQSNPTYYIRLANRDLVLRKKPPGTLLPSAHAI EREFRIMKALANAGVPVPNVLDLCEDSSVIGTPF YVMEYCPGLIYKDPSLPGLEPSHRRAIYTAMNTV LCKIHSVDLQAVGLEDYGKQGSTTWV/YSSRRA RGALLFLDWELSYPWGDPFADVGYSCLAHYLPS SFPVLRGINDCDLTQLGIPAAEEYFRMYCLQMGL PPTENWNFYMAFSFFRVAAILQGVYKRSLTGQA SSTYAEQTGKLTEFVSNLAWDFAVKEGFRVFKE MPFTNPLTRSYHTWARPQSQWCPTGSRSYSSVPE ASPAHTSRGGLVISPESLSPPVRELYHRLKHFME QRVYPAEPELQSHQASAARWSPSPLIEDLKVKQP W*GGRSGRTSWRLLALGCHT
3413	A	105	1573	PESRHQCFSDRSSHFLTMEMEQEKMTMNKELSP DAAAYCCSACHGDETWSYNHPIRGRAKSRSLSA SPALGSTKEFRRTRSLHGPCPVTTFGPKACVLQN PQTIMHIQDPASQRLTWNKSPKSVLVIKKMRDAS LLQPFKELCTHLMEENMIVYVEKKVLEDPAIASD ESFGAVKKKFCTFREDYDDISNQIDFIICLGGDGT LLYASSLFQGSVPPVMAFHLGSLGFLTPFSFENFQ SQVTQVIEGNAAVVL/RGSRLKVRVVKELRGKK TAVHNGLGEKGSQAAGLDMDVGKQAMQYQVL NEVVIDRGPSSYLSNVDVYLDGHLITTVQGD/G* GPQHLSWGP*AFLGRE*RLRLSLSGVIVSTPTGST AYAAAAGASMIHPNVPAIMITPICPHSLSFRPIVV PAGVELKIMLSPEARNTAWVSFDGRKRQEIRHG DSISITTSCYPLPSICVRDPVSDWFESLAQCLHWN VRKKQAHFEEEEEEEEEG
3414	A	20	2602	VIVNKNVNWINYIYYNQQQRAFHELKEKLMSAL ALGLPDLTKPFTFYESEREKMAVGVLTQTVGPW PRPVAYLSKQLDGVSKGWPPCLRALAATALLAQ EADKLTLGQNLNIKAPHAVVTLMNTKGHHWLT NARLTKYQSLPCENPHITIEVCNTLNPTTLLPVSE SPGEHNCVEVLDSVYSSRPDLRDQPWASSVDWE LYMDGSSFINSQGERCAGYAVVTLDAVIKAKLW LQGTSAQKAELIALTRAVELSEGQESLEELLGRY FYVSHLPAFAKAVAQLCITCRQHNARQSPTVSPH IQAYGAAPFEDLQVDFTEMPKCGGNKYLLVLTC TYSGWVEAYPTRTEKAYEVTRVLLRDLIPRFGLP LRIGSHNGPVFVADLDCVEINVDTGVIWATWIKN EKDPVQLQKGKSGPSCTKGQCNPLELVITNPLDP RWKKGERVTLGINGAGLNPRVNILVRGEVYKCS LEPVFQTFYDELNVPITEFPGKTRNLFLQLAEHV AQSLTVTSCYVCGGTVIADQWPWEARELVPTDP VPDEFPAQKNHPDNFWVLKASIIRQYYIARVEKD FTLPVGRLHGG/RSNHTEKNPFSKFPKLQTV*AHP ESHRDWTAPTGLYWICGHRAYTKLP\ASSCVIGTI KPSFFLLSIKTGELLGFPVYASR\KSIAIRN*NNDK WPPERIIQYYGPAT*AQDGSWGYRIPIYMINRIIRL QAVLKIITATGRALTILAQQETQMRNAIYQNRLA LDYLLAAEGEVCRKFNLTNCCLHIDNQGQVVED IVRDMTKVAHVPVQVWHGFDPGAMFRKWFPAL GGFKTLIIRVIIVIGTYLLLPRLLPVLLQMIKSFIAT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LVYQNASAQVYYINHY
3415	A	455	108	NMSWRGRSTYRPRPRRSLQPPELIGAMLEPTDEEPKEEKPPTKSRNPTPDQKREDDSG/SAA*DFKWPEPGKPIFQGAMVRPKTGG/CGCEGGY*CQGEDS\PKAEHFKMPEAGEGKSQV
3416	A	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWSCCLVQGGGDLVDVVQTNHGEDEAGGDTDSVDEARCKESQQEAQENLREDLCLESFAKDKILQIIEGSEREHEETRTKQAALDGEPLGGGQLTAVHLHPSKEQQGQEGGERQRGARTHHWRGWEKGRRVRLRPPSGKLRADQPVRKLGGPTPS/TELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPVRWKCPLPVEPRTRQLCRERTRKACPPKPRPPLGLPGDPTGPVTHHAPPVSPTGASGQERRAEPGAVSYAHASATK
3417	A	243	847	CLKYMYTYIFCPNCVSYKMKTDHFSLRYLHSSCAEDNKSSVDSSGQAAHPSKGKFFPHGTHWGTQCRGHISVLGWQCSCPSTGCRVGLGLAMCQTHAYIHTHTHTHTPTDYGAHHTDPLQRWGLGPR\KSEAGPLPQLSRDQSHPGPLSPGASPRSAGLPGWHPAHQEPRARGRCARDGLSLQTRLTNKYDIQCCQEMRK
3418	A	4073	1000	LDEYEARLTLANLDDFEEDNEDDDENRVNQEEKAAKITELINKLNFLDEAEKDLATVNSNPFDDPDAAELNPFGDPDSEEPITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEPEAFVTIKDSPPQSTKRKNIRPVDMSKYLYADSSKTEEEELDESNPFYEPKSTPPPNNLVNPVQELETERRVKRKAPAPPVLSPKTGVLNENTVSAGKDLSTSPKPSPIPSPVLGRKPNASQSLLVWCKEVTKNYRGVKITNFTTSWRNGLSFCAILHHFRPDLIDYKSLNPQDIKENNKKAYDGFASIGISRLLEPSDMVLLAIPDKLTVMTYLYQIRAHFSGQELNVVQIEENSSKSTYKVGNYETDTNSSVDQEKFYAELSDLKREPELQQPISGAVDFLSQDDSVFVNDSGVGESESEHQTPDDHLSPSTASPYCRRTKSDTEPQKSQQSSGRTSGSDDPGICSNTDSTQAQVLLGKKRLLKAETLELSDLYVSDKKKDMSPPFICEETDEQKLQTLDIGSNLEKEKLENSRSLECRSDPESPIKKTSLSPTSKLGYSYSRDLDLAKKKHASLRQTESDPDADRTTLNHADHSSKIVQHRLLSRQEELKERARVLLEQARRDAALKAGNKHNTNTATPFCNRQLSDQQDEERRRQLRERARQLIAEARSGVKMSELPSYGEMAAEKLKERSKASGDENDNIEIDTNEEIPEGFVVGGGDELTNLENDLDTPEQNSKLVDLKLKKLLEVQPQVANSPSSAAQKAVTESSEQDMKSGTEDLRTERLQKTTERFRNPVVFSKDSTVRKTQLQSFSQYIENRPEMKRQRSIQEDTKKGNEEKAAITETQRKPSEDEVLNKGFKDS\SQYVVGELAALENEQKQIDTRAALVEKRLRYLMDTGRNTEEEEAMMQEWFMLVNKKNALIRRMNQLSLLEKEHDLERRYELLNRELRAMLAIEDWQKTEAQKRREQLLLDELVALVNKRDALVRDLDAQEKQAEEEDEHLERTLEQNKGKMAKKEEKCVLQ
3419	A	4073	1000	LDEYEARLTLANLDDFEEDNEDDDENRVNQEEKAAKITELINKLNFLDEAEKDLATVNSNPFDDPDAAELNPFGDPDSEEPITETASPRKTEDSFYNNSYNP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				FKEVQTPQYLNPFDEPEAFVTIKDSPPQSTKRKNI RPVDMSKYLYADSSKTEEEELDESNPFYEPKSTP PPNNLVNPVQELETERRVKRKAPAPPVLSPKTGV LNENTVSAGKDLSTSPKPSPIPSPVLGRKPNASQS LLVWCKEVTKNYRGVKITNFTTSWRNGLSFCAI LHHFRPDLIDYKSLNPQDIKENNKKAYDGFASIGI SRLLEPSDMVLLAIPDKLTVMTYLYQIRAHFSGQ ELNVVQIEENSSKSTYKVGNYETDTNSSVDQEKF YAELSDLKREPELQQPISGAVDFLSQDDSVFVND SGVGESESEHQTPDDHLSPSTASPYCRRTKSDTEP QKSQQSSGRTSGSDDPGICSNTDSTQAQVLLGKK RLLKAETLELSDLYVSDKKKDMSPPFICEETDEQ KLQTLDIGSNLEKEKLENSRSLECRSDPESPIKKT SLSPTSKLGYSYSRDLDLAKKKHASLRQTESDPD ADRTTLNHADHSSKIVQHRLLSRQEELKERARVL LEQARRDAALKAGNKHNTNTATPFCNRQLSDQ QDEERRRQLRERARQLIAEARSGVKMSELPSYGE MAAEKLKERSKASGDENDNIEIDTNEEIPEGFVV GGGDELTNLENDLDTPEQNSKLVDLKLKKLLEV QPQVANSPSSAAQKAVTESSEQDMKSGTEDLRT ERLQKTTERFRNPVVFSKDSTVRKTQLQSFSQYI ENRPEMKRQRSIQEDTKKGNEEKAAITETQRKPS EDEVLNKGFKDS\SQYVVGELAALENEQKQIDTR AALVEKRLRYLMDTGRNTEEEEAMMQEWFML VNKKNALIRRMNQLSLLEKEHDLERRYELLNRE LRAMLAIEDWQKTEAQKRREQLLLDELVALVN KRDALVRDLDAQEKQAEEEDEHLERTLEQNKG KMAKKEEKCVLQ
3420	A	612	1058	ENLGPNYSHRLLHHPTFYKKIHKKHHEWTAPIG VISLYAHPIEHAVSNMLPVIVGPLVMGSHLSSITM WFSLALIITTISHCGYHLPFLPSPEFHDYHHLKFN QCYGVLGVLDHLHGTDTMFKQTKAYERHVLLL GFTPLSESIPDSPK
3421	A	23	2005	LLTPCDGRIPGRPSVGAESGSDFQQRRRRRRDPE EPEKTELSERELAVAVAVSQENDEENEERWVGP LPVEATLAKKRKVLEFERVYLDNLPSASMYERS YMHRDVITHVVCTKTDFIITASHDGHVKFWKKIE EGIEFVKHFRSHLGVIESIAVSSEGALFCSVGDDK AMKVFDVVNFDMINMLKLGYFPGQCEWIYCPG DAISSVAASEKSTGKIFIYDGRGDNQPLHIFDKLH TSPLTQIRLNPVYKAVVSSDKSGMIEYWTGPPHE YKFPKNVNWEYKTDTDLYEFAKCKAYPTSVCFS PDGKKIATIGSDRKVRIFRFVTGKLMRVFDESLS MFTELQQMRQQLPDMEFGRRMAVERELEKVDA VRLINIVFDETGHFVLYGTMLGIKVINVETNRCV RILGKQENIRVMQLALFQGIAKKHRAATTIEMKA SENPVLQNIQADPTIVCTSFKKNRFYMFTKREPE DTKSADSDRDVFNEKPSKEEVMAATQAEGPKRV SDSAIIHTSMGDIHTKLFPVECPKTVENFCVHSRN GYYNGHTFHRIIKGFMIQTGDPTGTGMGGESIWG GEFEDEFHSTLRHDRPYTLSMANAGSNTNGSQFF ITVVPTPWLDNKHTVFGRVTKGMEVVQRISN\VK VNPKTDKPYEDVSIINITVK
3422	A	2486	433	FVLVCAPLTWAGARHRRMAASKKPPRVRVNHQ DFQLRNLRIIEPNEVTHSGDTGVETDGRMPPKVT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SELLRQLRQAMRNSEYVTEPIQAYIIPSGDAHQSEYIAPCDCRRAFVSGFDGSAGTAIITEEHAAMWTDGRYFLQAAKQMDSNWTLMKMGLKDTPTQEDWLVSVLPEGSRVGVDPLIIPTDYWKKMAKVLRSAGHHLIPVKENLVDKIWTDRPERPCKPLLTLGLDYTGISWKDKVADLRLKMAERNVMWFVVTALDEIAWLFNLRGSDVEHNPVFFSYAIIGLETIMLFIDGDRIDAPSVKEHLLLDLGLEAEYRIQVHPYKSILSELKALCADLSPREKVWVSDKASYAVSETIPKDHRCCMPYTPICIAKA\VKNSA\ESEGMRRAHIKDAVALCELFNWLEKEVPKGGVTEISAADKAEEFRRQQADFVDLSFPTISSTGPNGAIIHYAPVPETNRTLSLDEVYLIDSGAQYKDGTTDVTRTMHFGTPTAYEKECFTYVLKGHIAVSAAVFPTGTKGHLLDSFARSALWDSGLDYLHGTGHGVGSFLNVHEGPCGISYKTFSDEPLEAGMIVTDEPGYYEDGAFGIRIENVVLVVPVKTKYNFNNRGSLTFEPLTLVPIQTKMIDVDSLTDKECDWLNNYHLTCRDVIGKELQKQGRQEALEWLIRETQPISKQH
3423	A	5515	934	FKMPENPATDKLQVLQVLDRLKMKLQEKGDTSQNEKLSMFYETLKSPLFNQILTLQQSIKQLKGQLNHIPSDCSANFDFSRKGLLVFTDGSITNGNVHRPSNNSTVSGLFPWTPKLGNEDFNSVIQQMAQGRQIEYIDIERPSTGGLGFSVVALRSQNLGKVDIFVKDVQPGSVADRDQRLKENDQILAINHTPLDQNISHQQAIALLQQTTGSLRLIVAREPVHTKSSTSSSLNDTTLPETVCWGHVEEVELINDGSGLGFGIVGGKTSGVVVRTIVPGGLADRDGRLQTGDHILKIGGTNVQGMTSEQVAQVLRNCGNSVRMLVARDPAGDISVTPPAPAALPVALPTVASKGPGSDSSLFETYNVELVRKDGQSLGIRIVGYVGTSHTGEASGIYVKSIIPGSAAYHNGHIQVNDKIVAVDGVNIQGFANHDVVEVLRNAGQVVHLTLVRRKTSSSTSPLEPPSDRGTVVEPLKPPALFLTGAVETETNVDGEDEEIKERIDTLKNDNIQALEKLEKVPDSPENELKSRWENLLGPDYEVMVATLDTQIADDAELQKYSKLLPIHTLRLGVEVDSFDGHHYISSIVSGGPVDTLGLLQPEDELLEVNGMQLYGKSRREAVSFLKEVPPPFTLVCCRRLFDDEASVDEPRRTETSLPETEVDHNMDVNTEEDDDGELALWSPEVKIVELVKDCKGLGFSILDYQDPLDPTRSVIVIRSLVADGVAERSGGLLPGDRLVSVNEYCLDNTSLAEAVEILKAVPPGLVHLGICKPLVEDNEEESCYILHSSSNEDKTEFSGTIHDINSSLILEAPKGFRDEPYFKEELVDEPFLDLGKSFHSQQKEIEQSKEAWEMHEFLTPRLQEMDEEREMLVDEEYELYQDPSPSMELYPLSHIQEATPVPSVNELHFGTQWLHDNEPSESQEARTGRTVYSQEAQPYGYCPENVMKENFVMESLPSVPSTEGNSQQGRFDDLENLNSLAKTSLDLGMIPNDVQGPSLLIDLPVVAQRREQEDLPLYQHQATRVISKASAYTGMLSSRYATDTCELPEREEGEGEETPNFSHWGPPRIVEIFREPNVSLGISIVGGQTVIKRLKNGEELKGIFIKQVLEDSPAGKTNALKTGDKILEVSGVDLQNASHSEAVEAIKNAGNPVVFIVQSLSSTPRVIPNVHNKANKITGNQNQDTQEKKEKRQGTAPPPMKLPPPYKALTDDSDENEEE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DAFTDQKIRQRYADLPGELHIIELEKDKNGLGLS LAGNKDRSRMSIFVVGINPEGPAAADGRMHIGD ELLEINNQILYGRSHQN\ASAIIKTAPSKVKLVFIR NEDAVNQMAVTPFPVPSSSPSSIEDQSGTEPISSEE \DGSLE\VGIKQLPESESFKLAVSQMKQQKYPTKV SFSSQEIPLAPASSYHSTDADFTGYGGFQAPLSVD PATCPIVPGQEMIIEISKRRSGLGLSIVGGKDTPLV NGVDLRNSSHEEAITALRQTPQKVRLVVYRDEA HYRDEENLEIFPVDLQKKAGRGLGLSIVGKR
3424	A	2223	1162	HASERVVQLPDFVWDQYTHSLGRVEREFKNRKR HTRRVKLVFDKGLPARPKSPLDPKKDGESLSYS MLPLSDGPEGSSSRPQMIRGRLCDDTKPETFNQL WTVEEQKKLEQLLIKYPPEEVESRRWQKIADELG NRTAKQVASRVQKYFIKLTKAGIPVPGRTPNLYI YSKKSSTSRRQHPLNKHLFKP\GTFMTSHEPPVY MDEDDDRSCFHSHMNTAVEDASDDESIPIMYRN LPEYKELLQFKKLKKQKLQHMQAESGFVQHVGF KCDNCGIEPIQG\VRW\HCR\DCPP\EMSL\DFC\DS C\SDCLHET\DIHKGDHQLEPIYRS\ETFLDRDYCV SQGTSYNYLDPNYFPANR
3425	A	2223	1162	HASERVVQLPDFVWDQYTHSLGRVEREFKNRKR HTRRVKLVFDKGLPARPKSPLDPKKDGESLSYS MLPLSDGPEGSSSRPQMIRGRLCDDTKPETFNQL WTVEEQKKLEQLLIKYPPEEVESRRWQKIADELG NRTAKQVASRVQKYFIKLTKAGIPVPGRTPNLYI YSKKSSTSRRQHPLNKHLFKP\GTFMTSHEPPVY MDEDDDRSCFHSHMNTAVEDASDDESIPIMYRN LPEYKELLQFKKLKKQKLQHMQAESGFVQHVGF KCDNCGIEPIQG\VRW\HCR\DCPP\EMSL\DFC\DS C\SDCLHET\DIHKGDHQLEPIYRS\ETFLDRDYCV SQGTSYNYLDPNYFPANR
3426	A	2	1553	LFVVVHDDPRWGTPRYWLGALYRNQQSSPTAPP GLLPLEYFPAAPHCSHSRQWRCSQTHRIHHHPQ MLGPCRQEICGITMAAGTLYTYPENWRAFKALI AAQYSGAQVRVLSAPPHFHFGQTNRTPEFLRKFP AGKVPAFEGDDGFCVFESNAIAYYVSNEELRGST PEAAAQVVQWVSFADSDIVPPASTWVFPTLGIM HHNKQATENAKEEVRRILGLLDAYLKTRTFLVG ERVTLADITVVCTLLWLYKQVLEPSFRQAFPNTN RWFLTCINQPQFRA\VFGEVKLCEKMAQF\DAKK FAETQPKKDTPRKEKGSREEKQKPQAERKEEKK AAAPAPEEEMDECEQALAAEPKAKDPFAHLPKS TFVLDEFKRKYSNEDTLSVALPYFWEHFDKDGW SLWYSEYRFPEELTQTFMSCNLITGMFQRLDKLR KNAFASVILFGTNNSSSISGVWVFRGQELAFPLSP DWQVDYESYTWRKLDPGSEETQTLVREYFSWE GAFQHVGKAFNQGKIFK
3427	A	755	52	TAARRRQKGTAARRRQKGTAARRRQKGTAARR RQKGTAARRRQKGTAARRRQKGTAARRRQKGT AARRRQKGTAARRRQKGTAARRRQKGTAARRR QKGLSNLDAAEWLPPKKG\GEKKKGPFLAINEV VT\REYPINILKRIHGVGFKKRAPRALKEIRKFAM KEMGTPDVRIDTRLNKAVWAKGIRNVPYRIRVR LSRKRNEDEDSPNKLYTLVTYVPVTTFKNLQTV NVDEN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3428	A	4	1939	LPLSLSFSEMPLPLLPMDLKGEPGPPGKPGPWGP PGPPGFPGKPGHGKPGLHGQPGPAGPPGFSRMG KAGPPGLPGNVGPPGQPGLRGEPGIRGDQGLRGP PGPPGLPGPSGITIPGKPGAQGVPGPPGFQGEPGP QGEPGPPGDRGLKGDNGVGQPGLPGAPGQGGAP GPPGLPGPAGLGKPGLDGLPGAPGDKGESGPPG VPGPRGEPGAVGPKGPPGVDGVGVPGAAGLPGP QGPSGAKGEPGTRGPPGLIGPTGYGMPGLPGPKG DRGPAGVPGLLGDRGEPGEDGEPGEQGPQGLGG PPGLPGSAGLPGRRGPPGPKGEAGPGGPPGVPGI RGDQGPSGLAGKPGVPGERGLPGAHGPPGPTGP KGEPGFTGRPGGPGVAGALGQKGDLGLPGQPGL RGPSGIPGLQGPAGPIGPQGLPGLKGEPGLPGPPG EGRAGEPGTAGP\RGPPGVPGSPGITGPPG\LPGPP GAPGAFDETGIAGLHLPNGGVEGAVLGKGGKPQ FGLGELSAHATPAFTAVLTSPLPASGMPVKFDRT LYNGHSGYNPATGIFTCPVGGVYYFAYHVHVKG TNVWVALYKNNVPATYTYDEYKKGYLDQASG GAVLQLRPNDQVWVQMPSDQANGLYSTEYIHSS FSGFLLCPT
3429	A	212	1075	EGLTGPCERVPFLLGRGPPHGATRAGHRRAVRW AGPESLPPLPRSLIMDSPRAGTHQGPLDAETEVG ADRCTSTAYQEQRPQVEQVGKQAPLSPGLPAMG GPGPGPCEDPAGAGGAGAGGSEPLVTVTVQCAF TVALRARRGADLSSLRALLGQALPHQ\AQLGQLS YLAPGEDGHWVPIPEEESLQRAWQDAAACPRGL QLQCRGAGGRPVLYQVVAQHSYSAQGPEDLGF RQGDTVDVLCEVDQAWLEGHCDGRIGIFPKCFV VPAGPRMSGAPGRLPRSQQGDQP
3430	A	799	1989	INKYINIRKKIKLLSPLPPLWSHLALLQASATKWV LTPAAFAGKLLSVFRQPLSSLWRSLVPLFCWLRA TFWLLATKRRKQQLVLRGPDETKEEEEDPPLPTT PTSVNYHFTRQCNYKCGFCFHTAKTSFVLPLEEA KRGLLLLK\EAG\LEKINFSGG\EPFLQDRGEYLGK LVRFCKVELRLPSVSI\VSNGSLIRERWFQNYG\E YLDILAISCDSFDEEVNCP\IGRGN\GKKNHVENL QKL\RRWCRDYRVPFKINSVINPF\NVEEDMTEQI KALNPVRWKVFQCLLIEGENCGEDA\LREAERFV IGDEEFERFLERHKEVSCLVPESNQKMKDSYLIL DEYMRFLNCRKGRKDPSKSILDVGVEEAIKFSGF DEKMFLKRGGKYIWSKADLKLDW
3431	A	5468	2146	ACGFLPGRCHFSTFKQCQEWLSRLSRATARPAKP EDLFAFAYHAWCLGLTEEDQHTHLCQPGEHIRC RQEAELARMGFDLQNVWRVSHINSNYKLCPSYP QKLLVPVWITDKELENVASFRSWKRIPVVVYRH LRNGAAIARCSQPEISWWGWRNADDEYLVTSIA KACALDPGTRATGGSLSTGNNDTSEACDADFDS SLTACSGVESTAAPQKLLILDARSYTAAVANRAK GGGCECEEYYPNCEVVFMGMANIHAIRNSFQYL RAVCSQMPDPSNWLSALESTKWLQHLSVMLKA AVLVANTVDREGRPVLVHCSDGWDRTPQIVALA KILLDPYYRTLEGFQVLVESDWLDFGHKFGDRC GHQENVEDQNEQCPVFLQWLDSVHQLLKQFPCL FEFNEAFLVKLVQHTYSCLYGTFLANNPC\EREK RNIYK/RGTCSVWALLRAGNKNFHNFLYTPSSD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MVLHPVCHVRALHLWTAVYLPASSPCTLGEEN MDLYLSPVAQSQEFSGRSLDRLPKTRSMDDLLS ACDTSSPLTRTSSDPNLNNHCQEVRVGLEPWHS NPEGSETSFVDSGVGGPQQTVGEVGLPPPLPSSQ KDYLSNKPFKSHKSCSPSYKLLNTAVPREMKSNT SDPEIKVLEETKGPAPDPSAQDELGRTLDGIGEPP EHCPETEAVSALSKVISNKCDGVCNFPESSQNSPT GTPQQAQPDSMLGVPSKCVLDHSLSTVCNPPSA ACQTPLDPSTDF\LNQDPSGSVASISHQEQLSSVP DLTHGEEDIGKRGNNRNGQLLENPRFGKMPLEL VRKPISQSQISEFSFLGSNWDSFQGMVTSFPSGEA TPRRLLSYGCCSKRPNSKQMRATGPCFGGQWAQ REGVKSPVCSSHSNGHCTGPGGKNQMWLSSHPK QVSSTKPVPLNCPSPVPPLYLDDDGLPFPTDVIQH RLRQIEAGYKQEVEQLRRQVRELQMRLDIRHCC APPAEPPMDYEDDFTCLKESDGSDTEDFGSDHSE DCLSEASWEPVDKKETEVTRWVPDHMASHCYN CDCEFWLAKRRHHCRNCGNVFCAGCCHLKLPIP DQQLYDPVLVCNSCYEHIQVSRARELMSQQLKK PIATASS
3432	A	36	1873	MTFFSSVADFIGLDPRIAAWLIDPSDATPSFEDLV EKYCEKSITVKVNSTYGNSSRNIVNQNVRENLKT LYRLTMDLCSKLKDYGLWQLFRTLELPLPILAV MESHAIQVNKEEMEKTSALLGARLKELEQEAHF VAGERFLITSNNQLREILFGKLKLHLLSQRNSLPR TGLQKYPSTVSEALNALRDLHPLPKIILEYRQVH KIKSTFVDGLLACMKKGSISSTWNQTGTVTGRLS AKHPNIQGISKHPIQITTPKNFKGKEDKILTISPRA MFVSSKGHTFLAADFSQIELRILTHLSGDPELLKL FQESERDDVFSTLTSQWKDVPVEQVTHADREQT KKVVYAVVYGAGKERLAACLGVPIQEAAQFLES FLQKYKKIKDFARAAIAQCHQTGCVVSIMGRRR PLPRIHAHDQQLRAQAERQAVNFVVQGSAADLC KLAMIHVFTAVAASHTLTARLVAQIHDELLFEVE DPQIPECAALVRRTMESLEQVPLKVSLSAGRSWG HLVPLQEAW\ALRQAHVALSLPATAWLPLGPLP APSPHPCIFRLHFVCSPRQQWEERTGFQQSIVWPS PRSPALYAPGRINPLGLGWPAIPWSKCLCKALKK K
3433	A	1481	476	IPPKERAPGIRASCLAITAGARPTSYGRVGCEGDV RLSPVSPLLAPPDPRLASRWEGRSRMKGKKGIVA ASGSETEDEDSMDIPLDLSSSAGSGKRRRRGNLP KESVQILRDWLYEHRYNAYPSEQEKALLSQQTH LSTLQVCNWFINARRRLLPDMLRKDGKDPNQFTI SRRGAKISETSSVESVMGIKNFMPALEETPFHSFT\ AGPNPTLG\RPLSAKP/SQSPGSVLARPSVICHTTV TAIERLSLSLSCQSVGCGQNT\DIQQIAT\RNLRDS SLMYPEDTCKSGPSTNTQSGLFNTPPPTPPDLNQ DFSGFQLLVDVALKRAAEMELQAKLTA
3434	A	1720	1243	NGPVPPGGSKTKWAGGSAAEGSPRLSPSPGAAQ VPALLRGEPRGGAAAGSFWKPLHQHSCGLRPPP/ PPD/RLSRLPGKTLSACDRENGARRPLLLGSTSFIP IGRRTYASAAEPVGSKAVLVTGCDSGFGFSLAKH LHSKGFLVFAGCLMKDKGHDGVKELDSLNSDRL RTVQLNVCSSEEVEKV/VGDCPLEPEGP\EKGMW

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GLVNNAGISTFGEVEFTSLETYKQVAEVNLWGTVRMTKSFLPLIRRAKGRVVNISSMLGRMANPARSPYCITKFGVEAFSDCLRYEMYPLGVKVSVVEPGNFIAATSLYSPESIQAIAKKMWEELPEVVRKDYGKKYFDEKIAKMETYCSSGSTDTSPVIDAVTHALTATTPYTRYHPMDYYWWLRMQIMTHLPGAISDMIYIR
3435	A	842	3595	ENQQQMLVAKEQRLHFLKQQERRQQQSISENEKLQKLKERVEAQENKLKKIRAMRGQVDYSKIMNGNLSAEIERFSAMFQEKKQEVQTAILRVDQLSQQLEDLKKGKLNGFQSYNGKLTGPAAVELKRLYQELQIRNQLNQEQNSKLQQQKELLNKRNMEVAMMDKRISELRERLYGKKIQACEKVFLNRVNGTSSPQSPLSTSGRVAAVGPYIQVPSAGSFPVLGDPIKPQSLSIASNAAHGRSKSANDGNWPTLKQNSSSSVKPVQVAGADWKDPSVEGSVKQGTVSSQPVPFSALGPTEKPGIEIGKVPPPIPGVGKQLPPSYGTYPSPTPLGPGSTSSLERRKEGSLPRPSAGLPSRQRPTLLPATGSTPQPGSSQQIQQRISVPPSPTYPPAGPPAFPAGDSKPELPLTVAIRPFLADKGSRPQSPRKGPQTVNSSSIYSMYLQQATPPKNYQPAAHSALNKSVKAVYGKPVLPSGSTSPSPLPFLHGSLSTGTPQPQPPSESTEKEPEQDGPAAPADGSTVESLPRPLSPTKLTPIVHSPLRYQSDADLEALRRKLANAPRPLKKRSSITEPEGPGGPNIQKLLYQRFNTLAGGMEGTPFYQPSPSQDFMVTLADVDNGNTNANGNLEELPPAQPTAPLPAEPAPSSDANDNELPSPEPEELICPQTTHQTAEPAEDNNNNVATVPTTEQIPSPVAEAPSPGEEQVPPAPLPPASHPPATSTNKRTNLKKPNSERTGHGLRVRFNPLALLLDASLEGEFDLVQRIIYEVEDPSKPNDEGITPLHNAVCAGHHHIVKFLLDFGVNVNAADSDGWTPLHCAASCNSVHLCKQLVESGAAIFASTISDIETAADKCEEMEEGYIQCSQFLYGVQEKLGVMNKGVAYALWDYEAQNSDELSFHEGDALTILRRKDE
3436	A	3	2604	GSTHASEKMKTGRSALVVTDTGDMSVLNSPRHQSCIMHVDMDCFFVSVGIRNRPDLKGKPVAVTSNRGTGRAPLRPGANPQLEWQYYQNKILKGKADIPDSSLWENPDSAQANGIDSVLSRAEIASCSYEARQLGIKNGMFFGHAKQLCPNLQAVPYDFHAYKEVAQTLYETLAS\YTHNIEAVSCDEALVDITEILAETKLTPDEFANAVRMEIKDQTKCAASVGIGSNILLARMATRKAKPDGQYHLKPEEVDDFIRGQLVTNLPGVGHSMESKLASLGIKTCGDLQYMTMAKLQKEFGPKTGQMLYRFCRGLDDRPVRTEKERKSVSAEINYGIRFTQPKEAEAFLLSLSEEIQRRLEATGMKGKRLTLKIMVRKPGAPVETAKFGGHGICDNIARTVTLDQATDNAKIIGKAMLNMFHTMKLNISDMRGVGIHVNQLVPTNLNPSTCPSRPSVQSSHFPSGSYSVRDVFQVQKAKKSTEEEHKEVFRAAVDLEISSASRTCTFLPPFPAHLPTSPDTNKAESSGKWNGLHTPVSVQSRLNLSIEVPSPSQLDQSVLEALPPDLREQVEQVCAVQQAESHGDKKKEPVNGCNTGILPQPVGTVLLQIPEPQESNSDAGINLIALPAFSQVDPEVFAALPAELQRELKAAYDQRQRQGENSTHQQSASASV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PKNPLLHLKAAVKEKKRNKKKKTIGSPKRIQSPL NNKLLNSPAKTLPGACGSPQKLIDGFLKHEGPPA EKPLEELSASTSGVPGLSSLQSDPAGCVRPPAPNL AGAVEFNDVKTLLREWITTISDPMEEDILQVVKY CTDLIEEKDLEKLDLVIKYMKRLMQQSVESVWN MAFDFILDNVQVVLQQTYGSTLKVT
3437	A	32	4038	SLLRLLKAQWGSSGAASEPVVLGEEGCGFPSTNE YPDLEEERATYPQEEDRFLTPGRAQLLWSPWSPL DQEEACASRQLHSLASFSTVTARRNPLHNPWGM ELAASENTDSPSPRPLRPGVTLPPGALTMNTKDT TEVAENSHHLKIFLPKKLLECLPRCPLLPPERLRW NTNEEIASYLITFEKHDEWLSCAPKTRPQNGSIIL YNRKKVKYRKDGYLWKKRKDGKTTREDHMKL KVQGMECLYGCYVHSSIVPTFHRRCYWLLQNPD IVLVHYLNVPALEDCGKGCSPIFCSISSDRREWLK WSREELLGQLKPMFHGIKWSCGNGTEEFSVEHL VQQILDTHPTKPAPRTHACLCSGGLGSGSLTHKC SSTKHRIISPKVEPRALTLTSIPHPHPPEPPPLIAPLP PELPKAHTSPSSSSSSSSGFAEPLEIRPSPPTSRGG SSRGGTAILLLTGLEQRAGGLTPTRHLAPQADPR PSMSLAVVVGTEPSAPPAPPSPAFDPDRFLNSPQR GQTYGGGQGVSPDFPEAEAAHTPCSALEPAAAL EPQAAARGPPPQSVAGGRRGNCFFIQDDDSGEEL KGHGAAPPIPSPPPSPPPSPAPLEPSSRVGRGEALF GGPVGASELEPFSLSSFPDLMGELISDEAPSIPAPT PQLSPALSTITDFSPEWSYPEGGVKVLITGPWTEA AEHYSCVFDHIAVPASLVQPGVLRCYCPAHEVG LVSLQVAGREGPLSASVLFEYRARRFLSLPSTQL DWLSLDDNQFRMSILERLEQMEKRMAEIAAAGQ VPCQGPDAPPVQDEGQGPGFEARVVVLVESMIP RSTWKGPERLAHGSPFRGMSLLHLAAAQGYARL IETLSQWRSVETGSLDLEQEVDPLNVDHFSCTPL MWACALGHLEAAVLLFRWNRQALSIPDSLGRLP LSVAHSRGHVRLARCLEELQRQEPSVEPPFALSP PSSSPDTGLSSVSSPSELSDGTFSVTSAYSSAPDGS PPPAPLPASEMTMEDMAPGQLSSGVPEAPLLLM DYEATNSKGPLSSLPALPPASDDGAAPEDADSPQ AVDVIPVDMISLAKQIIEATPERIKREDFVGLPEA GASMRERTGAVGLSETMSWLASYL\ENVDHFPS STPPSEL\PFER\GRLGLSLTAPSWAEFLSCIPPVGK IGKLIFALLTL\SD\QEQRELYEAARVIQTAFRKYK GRRLKEQQEVAAAVIQRCYRKYKQLTWIALKFA LYKKMTQAAILIQSKFRSYYEQKRFQQSRRAAV LIQQHYRSYRRRPGPPHRTSATLPARNKGSFLTK KQDQAARKIMRFLRRCRHRMRELKQNQELEGLP QPGLAT
3438	A	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQSLK VVTKRGSADGCTDWSIDIKKYQVLVGEPVRIKC ALFYGYIRTNYSLAQSAGLSLMWYKSSGPGDFE EPIAFDGSRMSKEEDSIWFRPTLLQDSGLYACVIR NSTYCMKVSISLTVGENDTGLCYNSKMKYFEKA ELSKSKEISCRDIEDFLLPTREPEILWYKECRTKT WRPSIVFKRDTLLIREVREDDIGNYTCELKYGGF VVRRTTELTVTAPLTDKPPKLLYPMESKLTIQET QLGDSANLTCRAFFGYSGDVSPLIYWMKGEKFIE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DLDENRVWESDI\KILKEHLGEQEVSISLIVDSVEE GDLGNYSCYVENGNGRRHASVLLHKRELMYTV ELAGGLGAILLLLVCLVTIYKCYKIEIMLFYRNHF GAEELDGDNKDYDAYLSYTKVDPDQWNQETGE EERFALEILPDMLEKHYGYKLFIPDRDLIPTGTYI EDVARCVDQSKRLIIVMTPNYVVRRGWSIFELET RLRNMLVTGEIKVILIECSELRGIMNYQEVEALK HTIKLLTVIKWHGPKCNKLNSKFWKRLQYEMPF KRIEPITHEQALDVSEQGPFGELQTVSAISMAAAT STALATAHPDLRSTFHNTYHSQMRQKHYYRSYE YDVPPTGTLPLTSIGNQHTYCNIPMTLINGQRPQT KSSREQNPDEAHTNSAILPLLPRETSISSVIW
3439	A	251	2037	GPGNSSILIGGGHLFLIRSCLNLLLLNSKENTEHT MAKKVAVIGAGVSGLSSIKCCVDEDLEPTCFERS DDIGGLWKFTERGSSLSVMIWPLALSLLRHGGFC YSDFPFHEDYPNFMNHEKFWDYLQEFAEHFDLL KYIQFKTTVCGITKRPDFSETGQWDVVTETEGKQ NRAVFDAVMVCTGHFLNPHLPLEAFPGIHKFKG QILHSQEYKIPEGFQGKRVLVIGLGNTGGDIAVEL SRTAAQVLLSTRTGTWVLGRSSDWGYPYNMMV TRRCCSFIAQVLPSRFLNWIQERKLNKRFNHEDY GLSITKGKKAKFIVNDELPNCILCGAITMKTSVIE FTETSAVFEDGTVEENIDVVIFTTGYTFSFPFFEEP LKSLCTKKIFLYKQVFPLNLERATLAIIGLIGLKGS ILSGTELQARWVTRVFKGLCKRPASQKLMMEAT EKEQLIKRGVFKDTSKDKFDYIAYMDDIAACIGT KPSIPLLFLKDPRLAWEVFFGPCTPYQYR\LMGPG KWDGARNAILTQWDRTLKPLKTRIVPDSSKAWP SM\SHYLKAWGAPVLLASLLLICK\SSLFLKLVRD KLQDRMSPYLVSLWRG
3440	A	1	3533	IMPCGSSRLLRGCWTHPNEPVSDLSYFDCIESVM ENSKVLGESMAGISQNAKTGDLPAFGECVGIASK ALCGLTEAAAQAAYLVGIFDPNSQAGHQGLVDP IQFARANQAIQMACQNLVDPGSSPSQVLSAATIV AKHTSALCNACRIASSKTANPVAKRHFVQSAKE VANSTANLVKTIKALDGDFSEDNRNKCRIATAPL IEAVENLTAFASNPEFVSIPAQISSEGSQAQEPILV SAKPMLESSSYLIRTARSLAINPKDPPTWSVLAG HSHTVSDSIKSLITSIRDKAPGQRECDYSIDGINRC IRDIEQASLAAVSQSLATRDDISVEALQEQLTSVV QEIGHLIDPIATAARGEAAQLGHKGTQLASYFEP LILAAVGVASKILDHQQQMTVLDQTKTLAESAL QMLYAAKEGGGNPKAQHTHDAITEAAQLMKEA VDDIMVTLNEAASEVGLVGGMVDAIAEAMSKL DEGTPPEPKGTFVDYQTTVVKYSKAIAVTAQEM MTKSVTNPEELGGLASQMTSDYGHLAFQGQMA AATAEPEEIGFQIRTRVQDLGHGCIFLVQKAG\AL QVCPTDSYTKRELIECARAVTEKVSLVLSALQAG NKGTQACITAATAVSGIIADLDTTIMFATAGTLN AENSETFADHRENILKTAKALVEDTKLLVSGAAS TPDKLAQAAQSSAATITQLAEVVKLGAASLGSD DPETQVVLINAIKDVAKALSDLISATKGAASKPV DDPSMYQLKGAAKVMVTNVTSLLKTVKAVEDE ATRGTRALEATIECIKQELTVFQSKDVPEKTSSPE ESIRMTKGITMATAKAVAAGNSCRQEDVIATAN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LSRKAVSDMLTACKQASFHPDVSDEVRTRALRFGTECTLGYLDLLEHVLVILQKPTPELKQQLAAFSKRVAGAVTELIQAAEAMKGTEWVDPEDPTVIAETELLGAAASIEAAAKKLEQLKPRAKPKQADETLDFEEQILEAAKSIAAATSALVKSASAAQRELVAQGKVGSIPANAADDGQWSQGLISAARMVAAATSSLCEAANASVQGHASEEKLISSAKQVAASTAQLLVACKVKADQDSEAMRRLQAAGNAVKRASDNLVRAAQKAAFGKADDDDVVVKTKFVGGIAQIIAAQEEMLKKERELEEARKKLAQIRQQQYKFLPTELREDEG
3441	A	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSALHSPAHRPPGFSVAQKPFGATYVWSSIINTLQTQVEVKKRRHRLKRHNDCFVGSEAVDVIFSHLIQNKYFGDVDIPRAKVVRVCQALMDYKVFEAVPTKVFGKDKKPTFEDSSCSLYRFTTIPNQDSQLGKENKLYSPARYADALFKSSDIRSASLEDLWENLSLKPANSPHVNISTTLSPQVINEVWQEETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPKRQSTMVNSSNYLDRGILKAYSDSQEDEWLSAAIDCLEYLPDQMVVEISRSFPEQPDRTDLVKELLFDAIGRYYSSREPLLNHLSDVHNGIAELLVNGKTEIALEATQLLLKLLDFQNREEFRRLLYFMAVAANPSEFKLQKESDNRMVVKRIFSKAIVDNKNLSKGKTDLLVLFL\MDHQKDVFKIPGTL\HKIVS\VK\LMAIQNGRDPNRDAGYIYCQRIDQRDYSNITEKTTIDELLYLLKTLDEDSKLSAKEKKK\LLGQFYKCHPDIFIEHFGD
3442	A	160	822	SPASGHCRLNGAAVAMFGCLVAGRLVQTAAQQVAEDKFVFDLPDYESINHVVVFMLGTIPFPEGMGGSVYFSYPDSNGMPVWQLLGFVTNGKPSAIFKISGLKSGEGSQHPFGAMNIVRTPSVAQIGISVELLDSMAQQTPVGNAAVSSVDSFTQFTQKMLDNFYNFASSFAVSQ/VPDDTQ/RPSEMFIPANVVLKWYENFQRRTSTEPSLLENIIWIKINF
3443	A	3	1373	SWHVRRRWLEATMAGGMKVAVSPAVGPGPWGSGVGGGGTVRLLLILSGCLVYGTAETDVNVVMLQESQVCEKRASQQFCYTNVLIPQWHDIWTRIQIRVNSSRLVRVTQVENEEKLKELEQFSIWNFFSSFLKEKLNDTYVNVGLYSTKTCLKVEIIEKDTKYSVIVIRRFDPKLFLVFLLGLMLFFCGDLLSRSQIFYYSTGMTVGIVASL\LIIIFILSKFMPKKSPIYVILVGGWSFSLYLIQLVFKNLQEIWRCYWQYLLSYVLTVGFMSFAVCYKYGPLENERSINLLTWTLQLMGLCFMYSGIQIPHIALAIIIALCTKNLEHPIQWLYITCRKVCKGAEKPVPPRLLTEEEYRIQGEVETRKALEELREFCNSPDCSAWKTVSRIQSPKRFADFVEGSSHLTPNEVSVHEQEYGLGSIIAQDEIYEEASSEEEDSYSRCPAITQNNFLT
3444	A	566	1718	KGLERTCCAMEESDSEKTTEKENLGPRMDPPLGEPG\GSLGWVLPNTAMKKKVLLMGKSGSGKTSMRSIIFANYIARDTRRLGATILDRIHSLQINSSLSTYSLVDSVGNTKTFDVEHSHVRFLGNLVLNLWDCGGQDTFMENYFTSQRDNIFRNVEVLIYVFDVESRELEKDMHYYQSCLEAILQNSPDAKIFCLVHKMDLVQEDQRDLIFKEREEDLRRLSRPLECSCFRTSIW

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DETLYKAWSSIVYQLIPNVQQLEMNLRNFAEIIE ADEVLLFERATFLVISHYQCKEQRDAHRFEKISNI IKQFKLSCSKLAASFQSMEVRNSNFAAFIDIFTSN TYVMVVMSDPSIPSAATLINIRNARKHFEKLERV DGPKQCLLMR
3445	A	566	1718	KGLERTCCAMEESDSEKTTEKENLGPRMDPPLG EPG\GSLGWVLPNTAMKKKVLLMGKSGSGKTS MRSIIFANYIARDTRRLGATILDRIHSLQINSSLST YSLVDSVGNTKTFDVEHSHVRFLGNLVLNLWDC GGQDTFMENYFTSQRDNIFRNVEVLIYVFDVESR ELEKDMHYYQSCLEAILQNSPDAKIFCLVHKMD LVQEDQRDLIFKEREEDLRRLSRPLECSCFRTSIW DETLYKAWSSIVYQLIPNVQQLEMNLRNFAEIIE ADEVLLFERATFLVISHYQCKEQRDAHRFEKISNI IKQFKLSCSKLAASFQSMEVRNSNFAAFIDIFTSN TYVMVVMSDPSIPSAATLINIRNARKHFEKLERV DGPKQCLLMR
3446	A	566	1718	KGLERTCCAMEESDSEKTTEKENLGPRMDPPLG EPG\GSLGWVLPNTAMKKKVLLMGKSGSGKTS MRSIIFANYIARDTRRLGATILDRIHSLQINSSLST YSLVDSVGNTKTFDVEHSHVRFLGNLVLNLWDC GGQDTFMENYFTSQRDNIFRNVEVLIYVFDVESR ELEKDMHYYQSCLEAILQNSPDAKIFCLVHKMD LVQEDQRDLIFKEREEDLRRLSRPLECSCFRTSIW DETLYKAWSSIVYQLIPNVQQLEMNLRNFAEIIE ADEVLLFERATFLVISHYQCKEQRDAHRFEKISNI IKQFKLSCSKLAASFQSMEVRNSNFAAFIDIFTSN TYVMVVMSDPSIPSAATLINIRNARKHFEKLERV DGPKQCLLMR
3447	A	1	2930	VLLGPLWDKLSTADHPVIVTMASKRKSTTPCMIP VKTVVLQDASMEAQPAETLPEGPQQDLPPEASA ASSEAAQNPSSTDGSTLANGHRSTLDGYLYSCK YCDFRSHDMTQFVGHMNSEHTDFNKDPTFVCSG CSFLAKTPEGLSLHNATCHSGEASFVWNVAKPD NHVVVEQSIPESTSTPDLAGEPSAEGADGQAEIIIT KTPIMKIMKGKAEAKKIHTLKENVPSQPVGEALP KLSTGEMEVREGDHSFINGAVPVRQASASSAKN PHAANGPLIGTVPVLPAGIAQFLSLQQQPPVHAQ HHVHQPLPTAKALPKVMIPLSSIPTYSAAMDSNS FLKNSFHKFPYPTKAELCYLTVVTKYPEEQLKIW FTAQRLKQGISWSPEEIEDARKKMFNTVIQSVPQ PTITVLNTPLVASAGNVQHLIQAALPGHVVGQPE GTGGGLLVTQPLMANGLQATSSPLPLTVTSVPK QPGVAPINTVCSNTTSAVKVVNAAQSLLTACPSI TSQAFLDASIYKNKKSHEQLSALKGSFCRNQFPG QSEVEHLTKVTGLSTREVRKWFSDRRYHCRNLK GSRAMIPGDHRSIIIDSVPEVSFSPSSKVPEVTCIPT TATLATHPSAKRQSWHQTPDFTPTKYKERAPEQ LRALESSFAQNPLPLDEELDRLRSETKMTRREIDS WFSERRKKVNAEETKKAEENASQEEEEAAEDEG GEEDLASELRVSGENGSLEMPSSHILAERKVSPIK INLKNLRVTEANGRNEIPGLGACDPEDDESNKLA EQLPGKVSCKKTAQQRHLLRQLFVQTQWPSNQD YDSIMAQTGLPRPEVVRWFGDSRYALKNGQLK WYEDYKRGNFPPGLLVIAPGNRELLQDYYMTHK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MLYEEDLQNLCDKTQMSSQQVKQWFAEKMGEETRAVADTGSEDQGPGTGELTAVHKGMGDTYSEVSENSESWEPRVPEASSEPFD\TSSPQAGRQLETD
3448	A	2	1324	FVARAEKGFRTREAHLLQVAGVGTGLQNGASLSGLASGVMAQRAFPNPYADYNKSLAEGYFDAAGRLTPEFSQRLTNKIRELLQQMERGLKSADPRDGTGYTGWAGIAVLYLHLYDVFGDPAYLQLAHGYVKQSLNCLTKRSITFLCGDAGPLAVAAVLYHKMNNEKQAEDCITRLIHLNKIDPHAPNEMLYGRIGYIYALLFVNKNFGVEKIPQSHIQQICETILTSGENLARKRNFTAKSPLMYEWYQEYYVGAAHGLAGIYYYLMQPSLQVSQGKLHSLVKPSVDYVCQLKFPSGNYPPCIGDNRDLLVHWCHGAPGVIYMLIQAYKVFR/EREKYLC\DAYQCADVIWQYGLLKKGYGLCY\GSAGNAYAFLTLYNLTQDMKYLYRACKFAEWCLEYGEHGCRTPDTPFSLFEGMAGTIYFL\ADLLFPTKAR\FPAFEL
3449	A	3	2389	SRHVTGAARSPSRAGPSDPPAMGDEDDDESCAVELRITEANLTGHEEKVSVENFELLKVLGTGAYGKVFLVRKAGGHDAGKLYAMKVLRKAALVQRAKTQEHTRTERSVLELVRQAPFLVTLHYAFQTDAKLHLILDYVSGGEMFTHLYQRQYFKEAEVRVYGGEIVLALEHLHKLGIIYRDLKLENVLLDSEGHIVLTDFGLSKEFLTEEKERTFSFCGTIEYMAPEIIRSKTGHGKAVDWWSLGILLFELLTGASPFTLEGERNTQAEVSRRILKCSPPFPPRIGPVAQDLLQRLLCKDPKKRLGAGPQGAQEVRNHPFFQGLDWVALAARKIPAPFRPQIRSELDVG\NFAEEFTRLEPVYSPPGQ\PPPGDPRIFQGYSFVAPSILFDHNNAVMTDGLEAPGAGDRPGRAAVARSAMMQDSPFFQQYELDLREPALGQGSFSVCRRCRQRQSGQEFAVKILSRRLEANTQREVAALRLCQSHPNVVNLHEVHHDQLHTYLVLELLRGGELLEHIRKKRHFSESEASQILRSLVSAVSFMHEEAGVVHRDLKPENILYADDTPGAPVKIIDFG/FSPRLRPQSPGVPMQTPSFTLQYAAPELLAQQGYDESCDLWSLGVILY\MMLSGQAPFQGASGQGGQSQAAEIMCKIREGRFSLDGEAWQGVSEEAKELVRGLLTVDPAKRLKLEGLRGSSWLQDGSARSSPPLRTPDVLESSGPAVRSGLNATFMAFNRGKREGFFLKSVENAPLAKRRKQKLRSATASRRGSPAPANPGRAPVASKGAPRRANGPLPPS
3450	A	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLRDSEDRSDSRAAQPAHDSGHGDDESPSTSSGTAGTSSVPELPGFYFDPEKKRYFRLLPGHNNCNPLTKESIRQKEMESKRLRLLQEEDRRKKIARMGFNASSMLRKSQLGFLNVTNYCHLAHELRLSCMERKKVQIRSMDPSALASDRFNLILADTNSDRLFTVNDVTVGGSKYGIINLQSLKTPTLKVFMHENLYFTNRKV\NSVCWASLNHLDSHILLCLMGLAETPGCATLLPASLFVNSHPAGIDRPG\MLCSFRIPGAWSCAWSLNIQANNCFSTGLSRRVLLTNVVTGHRQSFGTNSDVLAQQFALMAPLLFNGCRSGEIFAIDLRCGNQGKGWKATRLFHDSAVTSVRILQDEQYLMASDMAGKIKLWDLRTTKCVRQYEGHVNEYAYLPLHVHEEEGILVAVGQDCYTRIWSLHDARLLRTIPSPYPASKAD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				IPSVAFSSRLGGSRGAPGLLMAVGQDLYCYSYS
3451	A	19	6033	LLSAMLSHGAGLALWITLSLLQTGLAEPERCNFT LAESKASSHSVSIQWRILGSPCNFSLIYSSDTLGA ALCPTFRIDNTTYGCNLQDLQAGTIYNFKIISLDE ERTVVLQTDPLPPARFGVSKEKTTSTGLHVWWT PSSGKVTSYEVQLFDENNQKIQGVQIQESTSWNE YTFFNLTAGSKYNIAITAVSGGKRSFSVYTNGST VPSPVKDIGISTKANSLLISWSHGSGNVERYRLM LMDKGILVHGGVVDKHATSYAFHGLSPGYLYNL TVMTEAAGLQNYRWKLVRTAPMEVSNLKVTND GSLTSLKVKWQRPPG\NVDSYNITLSHKGTIKESR VLAPWIT\ETHFKELVPGRLY\QVTCSAVSLGELS AQKM\AVGRTFPDKVANLEANNNGRMRSLVVS WSPPAGDWEQYRILLFNDSVVLLNITVGKEETQ YVMDGTGLVPGRQYEVEVIVESGNLKNSERCQG RTVPLAVLQLRVKHANETSLSIMWQTPVAEWEK YIISLADRDLLLIHKSLSKDAKEFTFTDLVPGRKY MATVTSISGDLKNSSSVKGRTVPAQVTDLHVAN QGMTSSLFTNWTQAQGDVEFYQVLLIHENVVIK NESISSETSRYSFHSLKSGSLYSVVVTTVSGGISSR QVVVEGRTVPSSVSGVTVNNSGRNDYLSVSWLL APGDVDNYEVTLSHDGKVVQSLVIAKSVRECSF SSLTPGRLYTVTITTRSGKYENHSFSQERTVPDKV QGVSVSNSARSDYLRVSWVHATGDFDHYEVTIK NKNNFIQTKSIPKSENECVFVQLVPGRLYSVTVT TKSGQYEANEQGNGRTIPEPVKDLTLRNRSTEDL HVTWSGANGDVDQYEIQLLFNDMKVFPPFHLVN TATEYRFTSLTPGRQYKILVLTISGDVQQSAFIEG FTVPSAVKNIHISPNGATDSLTVNWTPGGGDVDS YTVSAFRHSQKVDSQTIPKHVFEHTFHRLEAGEQ YQIMIASVSGSLKNQINVVGRTVPASVQGVIADN AYSSYSLIVSWQKAAGVAERYDILLLTENGILLR NTSEPATTKQHKFEDLTPGKKYKIQILTVSGGLFS KEAQTEGRTVPAAVTDLRITENSTRHLSFRWTAS EGELSWYNIFLYNPDGNLQERAQVDPLVQSFSFQ NLLQGRMYKMVIVTHSGELSNESFIFGRTVPASV SHLRGSNRNTTDSLWFNWSPASGDFDFYELILYN PNGTKKENWKDKDLTEWRFQGLVPGRKYVLW VVTHSGDLSNKVTAESRTAPSPPSLMSFADIANT SLAITWKGPPDWTDYNDFELQWLPRDALTVFNP YNNRKSEGRIVYGLRPGRSYQFNVKTVSGDSWK TYSKPIFGSVRTKPDKIQNLHCRPQNSTAIACSWI PPDSDFDGYSIECRKMDTQEVEFSRKLEKEKSLL NIMMLVPHKRYLVSIKVQSAGMTSEVVEDSTIT MIDRPPPPPPHIRVNEKDVLISKSSINFTVNCSWFS DTNGAVKYFTVVVREADGSDELKPEQQHPLPSY LEYRHNASIRVYQTNYFASKCAENPNSNSKSFNI KLGAEMESLGGKCDPTQQKFCDGPLKPHTAYRI SIRAFTQLFDEDLKEFTKPLYSDTFFSLPITTESEP LFGAIEGVSAGLFLIGMLVAVVALLICRQKVSHG RERPSARLSIRRDRPLSVHLNLGQKGNRKTSCPIK INQFEGHFMKLQADSNYLLSKEYEELKDVGRNQ SCDIALLPENRGKNRYNNILPYDATRVKLSNVDD DPCSDYINASYIPGNNFRREYIVTQGPLPGTKDDF WKMVWEQNVHNIVMVTQCVEKGRVKCDHYW

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PADQDSLYYGDLILQMLSESVLPEWTIREFKICGE EQLDAHRLIRHFHYTVWPDHGVPETTQSLIQFVR TVRDYINRSPGAGPTVVHCSAGVGRTGTFIALDR ILQQLDSKDSVDIYGAV\HDLRLHRVHMVQTEC QYVYLHQCVRDVLRARKLRSEQENPLFPIYENV NPEYHRDPVYSRH
3452	A	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEGKS/ ARNSQLRIVLVGKTGAGKSATGNSILGRKVFHSG TAAKSITKKCEKRSSSWKETELVVVDTPGIFDTE VPNAETSKEIIRCILLTSPGPHALLLVVPLGRYTEE EHKATEKILKMFGERARSFMILIFTRKDDLGDTN LHDYLREAPEDIQDLMDIFGDRYCALNNKATGA EQEAQRAQLLGLIQRVVRENKEGCYTNRMYQR AEEEIQKQTQAMQELHRVELEREKARIREEYEEK IRKLEDKVEQEKRKKQMEKKLAEQEAHYAVRQ QRARTEVESKDGILELIMTALQIASFILLRLFAED
3453	A	2674	514	GPITFLKKKAKMKDMPLRIHVLLGLAITTLVQAV DKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCND LGLLTFPARLPANTQILLLQTNNIAKIEYSTDFPV NLTGLDLSQNNLSSVTNINGKKMPQLLSVYLEEN KLTELPEKCLSELSNLQELYINHNLLSTISPGAFIG LHNLLRLHLNSNRLQMINSKWFDALPNLEILMIG ENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNAL VGLENLESISFYDNRLIKVPHVALQKVVNLKFLD LNKNPINRIRRGDFSNMLHLKELGINNMPELISID SLAVDNLPDLRKIEATNNPRLSYIHPNAFFRLPKL ESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQ NVRQVHFRDMMEICLPLIAPESFPSNLNVEAGSY VSFHCRATA\EPQPEIYWITPSGQKLLPNT\LTDKF YVHSEGTLDINGVTPKEGGLYTCIATNLVGADLK SVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSW KASSKILKSSVKWTAFVKTENSHAAQSARIPSDV KVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVT TKGLHPDQKEYEKNNTTTLMACLGGLLGIIGVIC LISCLSPEMNCDGGHSYVRNYLQKPTFALGELYP PLINLWEAGKEKSTSLKVKATVIGLPTNMS
3454	A	1844	244	ERYLFATYVAPSATLDIGLQQEKKKEIYMKIQPP FEDLFDTAEEYILLLLLEPWTKMVKSDQIAYKKV ELVEETRQLDSTYFRKLQALHKETFSKKAEDTTC EIGTGILSLSNVSKRTEYWDNVPAEYKHFKFSDL LNNKLEFEHFRQFLETHSSSMDLMCWTDIEQFRR ITYRDRNQRKAKSIYIKNKYLNKKYFFGPNSPAS LYQQNQVMHLSGGWGKILHEQLDAPVLVEIQK HVQNRLENVWLPLFLASEQFAARQKIKVQMKDI AEELLLQKAEKKIGVWKPVESKWISSSCKIIAFRK ALLNPVTSRQFQRFVALKGDLLENGLLFWQEVQ KYKDLCHSHCDESVIQKKITTIINCFINSSIPPALQI DIPVEQAQKIIEHRKELGPYVFREAQMTFLGVMF KFWPQFCEFRKNLTDENIMSVLERRQEYNKQKK KLAVL/QNDEKSGKDGIKQYANTSVPAIKTALLS DSFLGLQPYGRQPTWCYSKYIEALEQERILLKIQE ELEK\SCLQACNLSQILRLALQLCL
3455	A	228	3330	APTAQAMMSFGGADALLGAPFAPLHGGGSLHY ALARKGGAGGTRSAAGSSSGFHSWTRTSVSSVS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ASPSRFRGAGAASSTDSLDTLSNGPEGCMVAVA TSRSEKEQLQALNDRFAGYIDKVRQLEAHNRSLE GEAAALRQQQAGRSAMGELYEREVREMRGAVL RLGAARGQLRLEQEHLLEDIAHVRQRLDDEARQ REEAEAAARALARFAQEAEAARVDLQKKAQAL QEECGYLRRHHQEEVGELLGQIQGSGAAQAQM QAETRDALKCDVTSALREIRAQLEGHAVQSTLQ SEEWFRVRLDRLSEAAKVNTDAMRSAQEEITEY RRQLQARTTELEALKSTKDSLERQRSELEDRHQA DIASYQEAIQQLDAELRNTKWEMAAQLREYQDL LNVKMALDIEIAAYRKLLEGEECRIGFGPIPFSLP EGLPKIPSVSTHIKVKSEEKIKVVEKSEKETVIVEE QTEETQVTEEVTEEEDKEAKEEEGKEEEGGEEEE AEGGEEETKSPPAEEAASPEKEAKSPVKEEAKSP AEAKSPEKEEAKSPAEVKSPEKAKSPAKEEAKSP PE\AKSPEKDGKQNFQAEVKSPEKAKSPAKEEAK SPAEAKSPEKAKSPVKEEAKSPAEAKSPVKEEAK SPAEVKSPEKAKSPTKEE\AKSPEKAKSPEKAKSP EKEEAKSPEKAKSPVKAEAKSPEKAKSPVKAEA KSPEKAKSPVKEEAKSPEKAKSPVKEEAKSPEKA KSPVKEEAKTPEKAKSPVKEEAKSPEKAKSPEKA KTLDVKSPEAKTPAKEEARSPADKFPEKAKSPVK EEVKSPEKAKSPLKEDAKAPEKEIPKKEEVKSPV KEEEKPQEVKVKEPPKKAEEEKAPATPKTEEKK DSKKEEAPKKEAPKPKVEEKKEPAVEKPKESKV EAKKEEAEDKKKVPTPEKEAPAKVEVKEDAKPK EKTEVAKKEPDDAKAKEPSKPAEKKEAAPEKKD TKEEKAKKPEEKPKTEAKAKEDDKTLSKEPSKP KAEKAEKSSSTDQKDSKPPEKATEDKAAKGK
3456	A	258	1463	YLSFIPGHASKSAPMNGHCFAENGPSQKSSLPPLL IPPSENLGPHEEDQVVCGFKKLTVNGVCASTPPL TPIKNSPSLFPCAPLCERGSRPLPPLPISEALSLDDT DCEVEFLTSSDTDFLLEDSTLSDFKYDVPG\RRSF RGCGQINYAYFDTPAVSAADLSYVSDQNG\GVP DPNPPPPQTHRRLRRSHSGPAGSFNKPAIRISNCCI HRASPNSDEDKPEVPPRVPIPPRPVKPDYRRWSA EVTSSTYSDEDRPPKVPPREPLSPSNSRTPSPKSLP SYLNGVMPPTQSFAPDPKYVSSKALQRQNSEGS ASKVPCILPIIENGKKVSSTHYYLLPERPPYLDKY EKFFREAKKKNGGAQIQPLPADCGISSATEKPDS KTKMDLGGHVKRKHLSYVGTP
3457	A	2	4869	FILSSSSSASSEHFHHHYSFGNWWPGSFKGHRMS LPFYQRCHQHYDLSYRNKDVRSTVSHYQREKKR SAVYTQGSTAYSSRSSAAHRRESEAFRRASASSS QQQASQHALSSEVSRKAASAYDYGSSHGLTDSS LLLDDYSSKLSPKPKRAKHSLLSGEEKENLPSDY MVPIFSGRQKHVSGITDTEEERIKEAAAYIAQRNL LASEEGITTPKQSTASKQTTASKQSTASKQSTASK QSTASRQSTASRQSVVSKQATSALQQEETSEKKS RKVVIRGKAERLSLRKTLEETETYHAKLNEDHLL HAPEFIIKPRSHTVWEKENVKLHCSIAGWPEPRV TWYKNQVPINVHANPGKYIIESRYGMHTLEINAC DFEDTAQYRASAMNVKGELSAYASVVVKRYKG EFDETRFHAGASTMPLSFGVTPYGYASRFEIHFD DKFDVSFGREGETMSLGCRVVITPEIKHFQPEIQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				WYRNGVPLSPSKWVQTLWSGERATLTFSHLNKE DEGLYTIRVRMGEYYEQYSAYVFVRDADAEIEG APAAPLDVKCLEANKDYIIISWKQPAVDGGSPIL GYFIDKCEVGTDSWSQCNDTPVKFARFPVTGLIE GRSYIFRVRAVNKMGIGFPSRVSEPVAALDPAEK ARLKS/PPLSTLDWT\VIVTEEEPSEGIVPGPPTDLS VTEATRSYVVLSWKPPGQRGHEGIMYFVEKCEA GTENWQRVNTELPVKSPRFALFDLAEGKSYCFR VRCSNSAGVGEPSEATEVTVVGDKLDIPKAPGKI IPSRNTDTSVVVSWEESKDAKELVGYYIEANVA GSGKWEPCNNNPVKTHRFTCHGLVTGQSYIFRV RAVNAAGLSEYSQDSEAIEVKAAIAPPSPPCDITC LESFRDSMVLGWKQPDKIGGAEITGYYVNYREV IDGVPGKWREANVKAVSEEAYKISNLKENMVY QFQVAAMNMAGLGAPSAVSECFKCEEWTIAVP GPPHSLKCSEVRKDSLVLQWKPPVHSGRTPVTG YFVDLKEAKAKEDQWRGLNEAAIKNVYLKVRG LKEGVSYVFRVRAINQAGVGKPSDLAGPVVAET RPGTKEVVVNVDDDGVISLNFECDKMTPKSEFS WSKDYVSTEDSPRLEVESKGNKTKMTFKDLGM DDLGIYSCDVTDTDGIASSYLIDEEELKRLLALSH EHKFPTVPVKSELAVEILEKGQVRF\WMQAEKLS GNAKVNYIFNEKGIFEGPKYKMHIDRNTGIIEMF MEKLQDEDEGTYTFQLQDGKATNHSTVVLVGD VFKKLQKEAEFQRQEWIRKQGPHFVEYLSWEVT GECNVLLKCKVANIKKETHIVWYKDEREISVDE KHDFKDGICTLLITEFSKKDAGIYEVILKDDRGK DKSRLKLVDEAFKELMMEVCKKIALSATDLKIQ STAEGIQLYSFVTYYVEDLKVNWSHNGSAIRYSD RVKTGVTGEQIWLQINEPTPNDKGKYVMELFDG KTGHQKTVDLSGQAYDEAYAEFQRLKQAAIAEK NRARVLGGLPDVVTIQEGKALNLTCNVWGDPPP EVSWLKNEKALASDDHCNLKFEAGRTAYFTING VSTADSGKYGLVVKNKYGSETSDFTVSVFIPEEE ARMAALESLKGGKKAK
3458	A	3963	827	LSRSSSDNNTNTLGRNVMSTATSPLMGAQSFPNL TTPGTTSTVTMSTSSVTSSSNVATATTVLSVGQS LSNTLTTSLTSTSSESDTGQEAEYSLYDFLDSCRA STLLAELDDDEDLPEPDEEDDENEDDNQEDQEY EEVMILRRPSLQRRAGSRSDVTHHAVTSQLPQVP AGAGSRPIGEQEEEEYETKGGRRRTWDDDYVLK RQFSALVPAFDPRPGRTNVQQTTDLEIPPPGTPHS ELLEEVECTPSPRLALTLKVTGLGTTREVELPLTN FRSTIFYYVQKLLQLSCNGNVKSDKLRRIWEPTY TIMYREMKDSDKEKENGKMGCWSIEHVEQYLG TDELPKNDLITYLQKNADAAFLRHWKLTGTNKS IRKNRNCSQLIAAYWDLG\EHGTK\SGLNQGAIST LQSSDILNLTKEQPQAKAGNGQNSCGVEDVLQL LRILYIVASDPYSRISQEDGDEQPQFTFPPDEFTS/ KKITTKILQQIEEPLALASGALPDWCEQLTSKCPF LIPFETRQLYFTCTAFGASRAIVWLQNRREATVE RTRTTSSVRRDDPGEFRVGRLKHERVKVPRGESL MEWAENVMQIHADRKSVLEVEFLGEEGTGLGPT LEFYALVAAEFQRTDLGAWLCDDNFPDDESRHV DLGGGLKPPGYYVQRSCGLFTAPFPQDSDELERI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TKLFHFLGIFLAKCIQDNRLVDLPISKPFFKLMCM GDIKSNMSKLIYESRGDRDLHCTESQSEASTEEG HDSLSVGSFEEDSKSEFILDPPKPKPPAWFNGILT WEDFELVNPHRARFLKEIKDLAIKRRQILSNKGL SEDEKNTKLQELVLKNPSGSGPPLSIEDLGLNFQF CPSSRIYGFTAVDLKPSGEDEMITMDNAEEYVDL MFDFCMHTGIQKQMEAFRDGFNKVFPMEKLSSF SHEEVQMILCGNQSPSWAAEDIINYTEPKLGYTR DSPGFLRFVRVLCGMSSDERKAFLQFTTGCSTLP PGGLANLHPRLTVVRKVDATDASYPSVNTCVHY LKLPEYSSEEIMRERLLAATMEKGFHLN
3459	A	88	603	SCGPRGLASLGLGFSGRCDDQNKGRS\DGPEAQA EACSGERTYQELLVNQNPIAQPLASRRLTRKLYK CIKKAVKQKQIRRGVKEVQKFVNKGEKGIMVLA GDTLPIEVYCHLPVMCEDRNLPYVYIPSKTDLGA AAGSKRPTCVIMVKPHEEYQEAYDECLEEVQSL PLPL
3460	A	139	1997	QVTNMSDKSELKAELERKKQRLAQIREEKKRKE EERKKKETDQKKEAVAPVQEESDLEKKRREAEA LLQSMGLTPESPIVPPPMSPSSKSVSTPSEAGSQD SGDGAVGSRRGPIKLGMAKITQVDFPPREIVTYT KETQTPVMAQPKEDEEEDDDVVAPKPPIEPEEEK TLKKDEEN\DSKAPPHELTEEEKQQILHSEEFLSFF DHSTRIVERALSEQINIFFDYSGRDF/ENDKEGEIQ AGAKLSLNRQFF\DER\WSKASGWVSCLDWSSQ YP\ELLVASYNNNEDAPHEPDGVALVWNMKYK KTTPEYVFHCQSAVMSATFAKFHPNLVVGGTYS GQIVLWDNRSNKRTPVQRTPLSAAAHTHPVYCV NVVGTQNAHNLISISTDGKICSWSLDMLSHPQDS MELVHKQSKAVAVTSMSFPVGDVNNFVVGSEE GSVYTACRHGSKAGISEMFEGHQGPITGIHCHAA VGAVDFSHLYVTSSFDWTVKLWTTKNNKPLYSF EDNAGYVYDVMWSPTHPALFACVDGMGRLDL WNLNNDTEVPTASISVEGNPALNRVRWTHSGRE IAVGDSEGQIVIYDVGEQIAVPRNDEWARFGRTL AEINANRADAEEEAATRIPA
3461	A	139	1997	QVTNMSDKSELKAELERKKQRLAQIREEKKRKE EERKKKETDQKKEAVAPVQEESDLEKKRREAEA LLQSMGLTPESPIVPPPMSPSSKSVSTPSEAGSQD SGDGAVGSRRGPIKLGMAKITQVDFPPREIVTYT KETQTPVMAQPKEDEEEDDDVVAPKPPIEPEEEK TLKKDEEN\DSKAPPHELTEEEKQQILHSEEFLSFF DHSTRIVERALSEQINIFFDYSGRDF/ENDKEGEIQ AGAKLSLNRQFF\DER\WSKASGWVSCLDWSSQ YP\ELLVASYNNNEDAPHEPDGVALVWNMKYK KTTPEYVFHCQSAVMSATFAKFHPNLVVGGTYS GQIVLWDNRSNKRTPVQRTPLSAAAHTHPVYCV NVVGTQNAHNLISISTDGKICSWSLDMLSHPQDS MELVHKQSKAVAVTSMSFPVGDVNNFVVGSEE GSVYTACRHGSKAGISEMFEGHQGPITGIHCHAA VGAVDFSHLYVTSSFDWTVKLWTTKNNKPLYSF EDNAGYVYDVMWSPTHPALFACVDGMGRLDL WNLNNDTEVPTASISVEGNPALNRVRWTHSGRE IAVGDSEGQIVIYDVGEQIAVPRNDEWARFGRTL AEINANRADAEEEAATRIPA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3462	A	2	2643	TAPEFSRSTHASAHASVARVLRNREIAQLKKEQRRQEFQIRALESQKRQQEMVLRRKTQEVSALRRLAKPMSERVAGRAGLKPPMLDSGAEVSASTTSSEAESGARSVSSIVRQWNRKINHFLGDHPAPTVNGTRPARKKFQKKGASQSFSKAARLKWQSLERRIIDIVMQRMTIVNLEADMERLIKKREELFLLQEALRRKRERLQAESPEEEKGLQELAEEIEVLAANIDYINDGITDCQATIVQLEETKEELDSTDTSVVISSCSLAEARLLLDNFLKASIDKGLQVAQKEAQIRLLEGRLRQTDMAGSSQNHLLLDALREKAEAHPELQALIYNVQQENGYASTDEEISEFSEGSFSQSFTMKGSTSHDDFKFKSEPKLSAQMKAVSAECLGPPLDISTKNITKSLASLVEIKEDGVGFSVRDPYYRDRVSRTVSLPTRGSTFPRQSRATETSPLTRRKSYDRGQPIRSTDVGFTPPSSPPTRPRNDRNVFSRLTSNQSQGSALDKSDDSDSSL\SEVLRGIISPVGGAKGARTAPLQCVSMAEGHTKPILCLDATDELLFTGSKDRSCKMWNLVTGQEIAALKGHPNNVVSIKYCSHSGLVFSVSTSYIKVWDIRDSAKCIRTLTSSGQVISGDACAATSTRAITSAQGEHQINQIALSPSGTMLYAASGNAVRIWELSRFQPVGKLTGHIGPVMCLTVTQTASQHDLVVTGSKDHYVKMFELGECVTGTIGPTHNFEPPHYDGIECLAIQGDILFSGSRDNGIKKWDLDQQELIQQIPNAHKDWVCALAFIPGRPMLLSACRAGVIKVWNVDNFTPIGEIKGHDSPINAICTNAKHIFTASSGCRVKVWNYVPGLTPCLPRRVLAIKGRATTLP
3463	A	198	3146	SGEPRPEPGNMATCIGEKIEDFKVGNLLGKGSFAGVYRAESIHTGLEVAIKMIDKKAMYKAGMVQRVQNEVKIHCQLKHPSILELYNYFEDSNYVYLVLEMCHNGEMNRYLKNRVKPFSENEARHFMHQIITGMLYLHSHGILHRDLTLSNLLLTRNMNIKIADFGLATQLKMPHEKHYTLCGTPNYISPEIATRSAHGLESDVWSLGCMFYTLLIGRPPFDTDTVKNTLNKVVLADYEMPTFLSIEAKDLIHQLLRRNPADRLSLSSVLDHPFMSRNSSTKSKDLGTVEDSIDSGHATISTAITASSSTSISGSLFDKRRLLIGQPLPNKMTVFPKNKSSTDFSSSGDGNSFYTQWGNQETSNSGRGRVIQDAEERPHSRYLRRAYSSDRSGTSNSQSQAKTYTMERCHSAEMLSVSKRSGGGENEERYSPTDNNANIFNFFKEKTSSSSGSFERPDNNQALSNHLCPGKTPFPFADPTPQTETVQQWFGNLQINAHLRKTTEYDSISPNRDFQGHPDLQKDTSKNAWTDTKVKKNSDASDNAHSVKQQNTMKYMTALHSKPEIIQQECVFGSDPLSEQSKTRGMEPPWGYQNRTLRSITSPLVAHRLKPIRQKTKKAVVSILDSEEVCVELVKEYASQEYVKEVLQISSDGNTITIYYPNGG\RGFPLA\DRPPSPT\DNISR\YSF\DNLPEKYWRKYQYASRFVQLVRSKSPKITYFTRYAKCILMENSPGADFEVWFYDGVKIHKTEDFIQVIEKTGKSYTLKSESEVNSLKEEIKMYMDHANEGHRICLALESIISEEERKTRSAPFFPIIGRKPGSTSSPKALSPPPSVDSNYPTRDRASFNRMVMHSAASPTQAPILNPSMVTNEGLGLTTTASGTDISSNSLKDCLPKSAQLLKSVFVKNVGWATQ\LTSGAVWVQFNDGSQLVVQAGVSSISYTSPNGQ\TTR\YGENEKLPDYIKQKLQCLSSILLMFSNPTPNFH

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3464	A	14	348	AVRTVSGTSLGPRSHSRSPGRCHCFSAVTFSSPRL AASEAPDPMEEWDVPQMKKEVESLKYQLAFQR EMASKTIPELLKWIEDGIPKDPFLNPDLMKNNPW V\EKGKCTIL
3465	A	5537	405	VRKLDRERVGAWWRGAWARHPRQEAGEHAKR RKGHAETPRGRRKGRAGRSAAAVGELRPARRSL ETSRAAAAMAKDSPSPLGASPKKPGCSSPAAAV LENQRRELEKLRAELEAERAGWRAERRRFAARE RQLREEAERERRQLADRLRSKWEAQRSRELRQL QEEMQREREAEIRQLLRWKEAEQRQLQQLLHRE RDGVVRQARELQRQLAEELVNRGHCSRPGASEV SAAQCRCRLQEVLAQLRWQTDGEQAARIRYLQ AALEVERQLFLKYILAHFRGHPALSGSPDPQAVH SLEEPLPQTSSGSCHAPKPACQLGSLDSLSAEVG VRSRSLGLVSSACSSSPDGLLSTHASSLDCFAPAC SRSLDSTRSLPKASKSEERPSSPDTSTPGSRRLSPP PSPLPPPPPPSAHRKLSNPRGGEGSESQPCEVLTPS PPGLGHHELIKLNWLLAKALWVLARRCYTLQEE NKQLRRAGCPYQADEKVKRLKVKRAELTGLAR RLADRARELQETNLRAVSAPIPGESCAGLELCQV FARQRARDLSEQASAPLAKDKQIEELRQECHLLQ ARVASGPCSDLHTGRGGPCTQWLNVRDLDRLQ RESQREVLRLQRQLMLQQGNGGAWPEAGGQSA TCEEVRRQMLALERELDQRRRECQELGAQAAPA RRRGEEAETQLQAALLKNAWLAEENGRLQAKT DWVRKVEAENSEVRGHLGRACQERDASGLIAEQ LLQQAARGQDRQQQLQRDPQKALCDLHPSWKEI QALQCRPGHPPEQPWETSQMPESQVKGSRRPKF HARAEDYAVSQPNRDIQEKREASLEESPVALGES ASVPQVSETVPASQPLSKKTSSQSNSSSEGSMWA TVPSSPTLDRDTASEVDDLEPDSVSLALEMGGSA APAAPKLKIFMAQYNYNPFEGPNDHPEGELPLTA GDYIYIFGDMDEDGFYEGELEDGRRGLVPSNFVE QIPDSYIPGCLPAKSPDLGPSQLPAGQDEALEEDS LLSGKAQGVVDRGLCQMVRVGSKTEVATEILDT KTEACQLGLLQSMGKQGLSRPLLGTKGVLRMAP MQLHLQNVTATSANITWVYSSHRHPHVVYLDD REHALTPAGVSCYTFQGLCPGTHYRARVEVRLP RDLLQVYWGTMSSTVTFDTLLAGPPYPPLDVLV ERHASPGVLVVSWLPVTIDSAGSSNGVQVTGYA VYADGLKVCEVADATAGSTLLEFSQLQVPLTWQ KVSVRTMSLCGESLDSVPAQIPEDFFMCHRWPET PPFSYTCGDPSTYRVTFPVCPQKLSLAPPSAKASP HNPGSCGEPQAKFLEAFFEEPPRRQSPVSNLGSE GECPSSGAGSQAQELAEAWEGCRKDLLFQKSPQ NHRPPSVSDQTGEKENCYQHMGTSKSPAPGFIHL RTECGPRKEPCQEKAALERVLRQKQDAQGFTPP QLGASQQYASDFHNVLKEEQEALCLDLWGTERR EERREPEPHSRQGQALGVKRGCQLHEPSSALCPA PSAKVIKMPRGGPQQLGTGANTPARVFVALSDY NPLVMSANLKAAEEELVFQKRQLLRVWGSQDT HDFYLSECNRQVGNIPGRLVAEMEVGTEQTDRR WRSPAQGHLPSVAHLEDFQGLTIPQGSSLVLQGN SKRLPLWTPKIMIAALDYDPGDGQMGGQGKGRL ALRAGDVVMVY\GPMDDQGFYYGELGGHRG\L

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VPANLRIKMSSQGH
3466	A	1	1111	MSKPPDLLLRLLRGAPRQRVCTLFIIGFKFTFFVSIMIYWHVVGEPKEKGQLYNLPAEIPCPTLTPPTPPSHGPTPGNIFFLETSDRTNPNFLFMCSVESAARTHPESHVLVLMKGLPGGNASLPRHLGISLLSCFPNVQMLPLDLRELFRDTPLADWYAAVQGRWEPYLLPVLSDASRIALMWKFGGIYLDTDFIVLKNLRNLTNVLGTQSRYVLNGAFLAFERRHEFMALCMRDFVDHYNGWIWGHQGPQLLTRVFKKWCSIRSLAESRACRGVTTLPPEAFYPIPWQDWKKYFEDINPEELPRLLSATYAVHVWNKKSQGTRFEATSRALLAQLHARYCPTTHE/DHENVLVKGPAGHLPNLLMGHW
3467	A	1	2175	MAKVILKQSKQCKNLLTCKVAQVCPVCGCLHCYFWWLSGLESRRPSSPLIDIKPIEFGVLSAKKEPIQPSVLRRTYNPDDYFRKFEPHLYSLDSNSDDVDSLTDEEILSKYQLGMLHFSTQYDLLHNHLTVRVIEARDLPPPISHDGSRQDMAHSNPYVKICLLPDQKNSKQTGVKRKTQKPVFEERYTFEIPFLEAQRRTLLLTVVDFDKFSRHCVIGKVSVPLCEVDLVKGGHWWKAHDSQFSAPGLPADQQFFADLFSGLVLNPQLLGRVWFASQPASLPVGSLCIDFPRLDIVLRGEYGNLLEAKQQRLVEGEMLFIPARAANLPVNNKPVMLLSLVFAPTWLGLSFYDSRTTSLLHPARQIQLP\SLQRGEGEAMLS\ALTLFSRSPLEQNIIQPLVLSLLHLCGSVVNMPPGNSQPRGDFLYHSICTWVQDNYAQPLTRESVAQFFNITPNHLSKLFAQHGTMRFIEYVRWVRMAKARMILQKYHLSIHEVAQRCGFPDSDYFCRVFRRQFGMDYVDILQIHRWDYNTPIEETLEALNDVVKAGKARYIGASSMHASQFAQALELQKQHGWAQFVSMQDHYNLIYREEEREMLPLCYQEGVAVIPWSPLARGRLTRPWGETTARLVSDEVGKNLYKESDENDAQIAERLTGVSEELGATRAQVALAWLLSKPGIAAPIIGTSREEQLDELLNAVDITLKPEQIAELETPYKPHPVVGFK
3468	A	147	3209	ALPLPLPTLYPGMSRRKQRKPQQLISDCEGPSASENGDASEEDHPQVCAKCCAQFTDPTEFLAHQNACSTDPPVMVIIGGQENPNNSSASSEPRPEGHNNPQVMDTEHSNPPDSGSSVPTDPTWGPERRGEESSGHFLVAATGTAAGGGGGLILASPKLGATPLPPESTPAPPPPPPPPPPGVGSGHLNIPLILEELRVLQQRQIHQMQMTEQICRQVLLLGSLGQTVGAPASPSELPGTGTASSTKPLLPLFSPIKPVQTSKTLASSSSSSSSSGAETPKQAFFHLYHPLGSQHPFSAGGVGRSHKPTPAPSPALPGSTDQLIASPHLAFPSTTGLLAAQCLGAARGLEATASPGLLKPKNGSGELSYGEVMGPLEKPGGRHKCRFCAKVFGSDSALQIHLRSHTGERPYKCNVCGNRFTTRGNLKVHFHRHREKYPHVQMNPHPVPEHLDYVITSSGLPYGMSVPPEKAEEEAATPGGGVERKPLVASTTALSATESLTLLSTSAGTATAPGLPAFNKFVLMKAVEPKNKADENTPPGSEGSAISGVAESSTATRMQLSKLVTSLPSWALLTNHFKSTGSFPLPLCARALG\ASPSETSKLQQLVEKIDRQGAVAVTSAASGAPTTSAPAPSSSASSGPNQCVICLRVLSCPRALRLHYGQHGGERPFKCKVCGRAFSTRGNLRAHFVGHKASPAARAQNSCPICQKKFT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NAVTLQQHVRMHLGGQIPNGGTALPEGGGAAQ ENGSEQSTVSGAGSFPQQQSQQPSPEEELSEEEEE EDEEEEEDVTDEDSLAGRGSESGGEKAISVRGDS EEASGAEEEVGTVAAAATAGKEMDSNEKTTQQS SLPPPPPPDSLDQPQPMEQGSSGVLGGKEEGGKP ERSSSPASALTPEGEATSVTLVEELSLQEAMRKEP GESSSRKACEVCGQAFPSQAAL\EEH\QKTHPKEG PLF\TCVFCRQGFLERATLKKHMLLAHHQVQPFA PHGPQNIAALSLVPGCSPSITSTGLSPFPRKDDPTI P
3469	A	3	5664	NLRPLSFALFLGDPNMANLEESFPRGGTRKIHKP EKAFQQSVEQDNLFDISTEEGSTKRKKSQKGPAK TKKLKIEKRESSKSAREKFEILSVESLCEGMRILG CVKEVNELELVISLPNGLQGFVQVTEICDAYTKK LNEQVTQEQPLKDLLHLPELFSPGMLVRCVVSSL GITDRGKKSVKLSLNPKNVNRVLSAEALKPGML LTGTVSSLEDHGYLVDIGVDGTRAFLPLLKAQEY IRQKNKGAKLKVGQYLNCIVEKVKGNGGVVSLS VGHSEVSTAIATEQQSWNLNNLLPGLVVKAQVQ KVTPFGLTLNFLTFFTGVVDFMHLDPKKAGTYFS NQAVRACILCVHPRTRVVHLSLRPIFLQPGRPLTR LSCQNLGAVLDDVPVQGFFKKAGATFRLKDGVL AYARLSHLSDSKNVFNPEAFKPGNTHKCRIIDYS QMDELALLSLRTSIIEAQYLRYHDIEPGAVVKGT VLTIKSYGMLVKVGEQMRGLVPPMHLADILMK NPEKKYHIGDEVKCRVLLCDPEAKKLMMTLKKT LIESKLPVITCYADAKPGLQTHGFIIRVKDYGCIV KFYNNVQGLVPKHELSTEYIPDPERVFYTGQVV KVVVLNCEPSKERMLLSFKLSSDPEPKKEPAGHS QKKGKAINIGQLVDVKVLEKTKDGLEVAVLPHN IRAFLPTSHLSDHVANGPLLHHWLQAGDILHRVL CLSQSEGRVLLCRKPALVSTVEGGQDPKNFSEIH PGMLLIGFVKSIKDYGVFIQLPSGLSGLAPKAIMS DKFVTSTSDHFVEGQTVAAKVTNVDEEKQRMLL SLRLSDCGLGDLAITSLLLLNQCLEELQGVRSLM SNRDSVLIQTLAEMTPGMFLDLVVQEVLEDGSV VFSGGPVPDLVLKASRYHRAGQEVESGQKKKVV ILNVDLLKLEVHVSLHQ\DLV\NRKARKLRKGSE HQAIVQHLEKSFAIASLVETGHLAAFSLTSHLND TFRFDSEKLQVGQGVSLTLKTTEPGVTGLLLAVE GPAAKRTMRPTQKDSETVDEDEEVDPALTVGTI KKHTLSIGDMVTGTVKSIKPTHVVVTLEDGIIGCI HASHILDDVPEGTSPTTKLKVGKTVTARVIGGRD MKTFKYLPISHPRFVRTIPELSVRPSELEDGHTAL NTHSVSPMEKIKQYQAGQTVTCFLKKYNVVKK WLEVEIAPDIRGRIPLLLTSLSFKVLKHPDKKFRV GQALRATVVGPDSSKTFLCLSLTGPHKLEEGEVA MGRVVKVTPNEGLTVSFPFGKIGTVSIFHMSDSY SETPLEDFVPQKVVRCYILSTADNVLTLSLRSSRT NPETKSKVEDPEINSIQDIKEGQLLRGYVGSIQPH GVFFRLGPSVVGLARYSHVSQHSPSKKALYNKH LPEGKLLTARVLRLNHQKNLVELSFLPGDTGKPD VLSASLEGQLTKQEERKTEAEERDQKGEKKNQK RNEKKNQKGQEEVEMPSKEKQQPQKPQAQKRG GRECRESGSEQERVSKKPKKAGLSEEDDSLVDV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YYREGKEEAEETNVLPKEKQTKPAEAPRLQLSSGFAWNVGLDSLTPALPPLAESSDSEEDEKPHQATIKKSKKERELEKQKAEKELSRTEEALMDPGRQPESADDFDRLVLSSPNSSILWLQYMAFHLQATEIEKARAVAERALKTISFREEQEKLNVWVALLNLENMYGSQESLTKVFERAVQYNEPLKVFLHLADIYAKSEKFQEAGELYNRMLKRFRQEKAVWIKYGAFLLRRSQAAASHRVLQRALECLPSKEHVDVIAKFAQLEFQLGDAERAKAIFENTLSTYPKRTDVWSVYIDMTIKHGSQKDVRDIFERVIHLSLAPKRMKFFFKRYLDYEKQHGTEKDVQAVKAKALEYVEAKSSVLED
3470	A	2334	1226	TAAAPVAPGTMDDATVLRKKGYIVGINLGKGSYAKVKSAYSERLKFNVAVKIIARKKTPTDFVERFLPREMDILATVNHGSIIKTYEIFETSDGRIYIIMELGVQGDLLEFIKCQGALHEDVARKMFRQLSSAVKYCHDLDIVHRDLKCENLLLDKDFNIKLSDFGFSKRCLRDSNGRIILSKTFCGSAAYAAPEVLQSIPYQPKVYDIWSLGVILYIMVCGSMPYDDSDIRKMLRIQKEHRVDFPRSKNLTCECKDLIYRMLQ\PDVS\KRLHIDEILSHSWLQPPKPK\ATSSASFKREGEGKYRAECKLDTKTGLRPDHRPDHKLGAKTQHRLLVVPENENRMEDRLAETSRAKDHHISGAEVGKAST
3471	A	537	148	TERGAPQHPTLPLPSLTPSSVHTGQPKTTPSVILFLPSCEEPQANKATLVCLMNN/FYPGILMVTWKADGTLITQSVEKTTPSKQSNNKYVASSYLSLTPEQWRSRRSYSCQVMQEGSTVEKSVAPAECS
3472	A	1	2272	DKPTRHKTYLSSSWAKMAAAEGPVGDGELWQTWLPNHVVFLRLREGLKNQSPTEAEKPASSSLPSSPPPQLLTRNVVFGLGGELFLWDGEDSSFLVVRLRGPSGGGEEPALSQYQRLLCINPPLFEIYQVLLSPTQHHVALIGIKGLMVLELPKRWGKNSEFEGGKSTVNCSTTPVAERFFTSSTSLTLKHAAWYPSEILDPHVVLLTSDNVIRIYSLREPQTPTNVIILSEAEEESLVLNKGRAYTASLGETAVAFDFGPLAAVPKTLFGQNGKDEVVAYPLYILYENGETFLTYISLLHSPGN/IWKAVGSIAHAS\AAEDNYGYDACAVLCLPCVPNILVIATESGMLYHCVVLEGEEEDDHTSEKSWDSRIDLIPSLYVFECVELELALKLASGEDDPFDSDFSCPVKLHRDPKCPSRYHCTHEAGVHSVGLTWIHKLHKFLGSDEEDKDSLQELSTEQKCFVEHILCTKPLPCRQPAPIRGFWIVPDILGPTMICITSTYECLIWPLLSTVHPASPPLLCTREDVEVAESPLRVLAETPDSFEKHIRSILQRSVANPAFLKASEKDIAPPPEECLQLLSRATQVFREQYILKQDLAKEEIQRRVKLLCDQKKKQLEDLSYCREERKSLREMAERLADKYEEAKEKQEDIMNRMKKLLHSFHSELPVLSDSERDMKKELQLIPDQLRHLGNAIKQVTMKKDYQQQKMEKVLSLPKPTIILSAYQRKCIQSILKEEGEHIREMVKQINDIRNHVNF
3473	A	1	2272	DKPTRHKTYLSSSWAKMAAAEGPVGDGELWQTWLPNHVVFLRLREGLKNQSPTEAEKPASSSLPSSPPPQLLTRNVVFGLGGELFLWDGEDSSFLVVRLRGPSGGGEEPALSQYQRLLCINPPLFEIYQVLLSPTQHHVALIGIKGLMVLELPKRWGKNSEFEGGKST

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VNCSTTPVAERFFTSSTSLTLKHAAWYPSEILDPH VVLLTSDNVIRIYSLREPQTPTNVIILSEAEEESLV LNKGRAYTASLGETAVAFDFGPLAAVPKTLFGQ NGKDEVVAYPLYILYENGETFLTYISLLHSPGN/I WKAVGSIAHAS\AAEDNYGYDACAVLCLPCVPN ILVIATESGMLYHCVVLEGEEEDDHTSEKSWDSR IDLIPSLYVFECVELELALKLASGEDDPFDSDFSC PVKLHRDPKCPSRYHCTHEAGVHSVGLTWIHKL HKFLGSDEEDKDSLQELSTEQKCFVEHILCTKPLP CRQPAPIRGFWIVPDILGPTMICITSTYECLIWPLL STVHPASPPLLCTREDVEVAESPLRVLAETPDSFE KHIRSILQRSVANPAFLKASEKDIAPPPEECLQLLS RATQVFREQYILKQDLAKEEIQRRVKLLCDQKK KQLEDLSYCREERKSLREMAERLADKYEEAKEK QEDIMNRMKKLLHSFHSELPVLSDSERDMKKEL QLIPDQLRHLGNAIKQVTMKKDYQQQKMEKVL SLPKPTIILSAYQRKCIQSILKEEGEHIREMVKQIN DIRNHVNF
3474	A	4344	2550	DRRREPERHVRVKQRTSVLNMLRRLDKIRFRGH KRDDFLDLAESPNASDTECSDEIPLKVPRTSPRDS EELRDPAGPGTLIMATGVQDFNRTEFDRLNEIKG HLEIALLEKHFLQEELRKLREETNAEMLRQELDR ERQRRMELEQKVQEVLKARTEEQMAQQPPKGQ AQASNGAERRSQGLSSRLQKWFYERFGEYVEDF RFQPEENTVETEEPLSARRLTENMRRLKRGAKPV TNFVKNLSALSDWYSVYTSAIAFTVYMNAVWH GWAIPLFLFLAILRLSLNYLIARGWRIQWSIVPEV SEPVEPPKEDLTVSEKFQLVLDVAQKAQNLFGK MADILEKIKNLFMWVQPEITQKLYVALWAAFLA SCFFPYRLVGLAVGLYAGIKFFLIDFIFKRCPRLR AKYDTPYIIWRSLPTDPQLKERSSAAVSRRLQTTS SRSYVPSAPAGLGKEEDAGRFHSTKKGNFHEIFN LTENERPLAVCENGWRCCLINRDRKMPTDYIRN GVLYVT\ENYLCFESSKSGSSKRNKVIKLVDITDI QKYKVLSVLPGSGMGIAVSTPSTQKPLVFGAMV HRDEAFETILSQYIKITSAAASGGDS
3475	A	2	1126	TAARRRQKGAAAAAETHGQAKAKSGWLKPYYF IELMESRKDITNQEELWKMKPRRNLEEDDYLHK DTGETSMLKRPVLLHLHQTAHADEFDCPSELQH TQELFPQWHLPIKIAAIIASLTFLYTLLREVIHPLA TSHQQYFYKIPILVINKVLPMVSITLLALVYLPGV IAAIVQLHNGTKYKKFPHWLDKWMLTRKQFGL LSFFFAVLHAIYSLSYPMRRSYRYKLLNWAYQQ VQQNKEDAL\IEHDVWRMEIYVSLGIVGLAILAL LAVTSIPSVSDSLTWREFHYIQSKLGIVSLLLGTIH ALIFAWNKWIDIKQFVWYTPPTFMIAVFLPIVVLI FKSILFLPCLRKKILKIRHGWEDVTKINKTEICSQL
3476	A	143	3191	AKAPPTGESSEPEAKVLHTKRLYRAVVEAVHRL DLILCNKTAYQEVFKPENISLRNKLRELCVKLMF LHPVDYGRKAEELLWRKVYYEVIQLIKTNKKHI HSRSTLECAYRTHLVAGIGFYQHLLLYIQSHYQL ELQCCIDWTHVTDPLIGCKKPVSASGKEMDWAQ MACHRCLVYLGDLSRYQNELAGVDTELLAERFY YQALSVAPQIGMPFNQLGTLAGSKYYNVEAMY CYLRCIQSEVSFEGAYGNLKRLYDKAAKMYHQL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KKCETRKLSPGKKRCKDIKRLLVNFMYLQSLLQPKSSSVDSELTSLCQSVLEDFNLCLFYLPSSPNLSLASEDEEEYESGYAFLPDLLIFQMVIICLMCVHSLERAGSKQYSAAIAFTLALFSHLVNHVNIRLQAELEEGENPVPAFQSDGTDEPESKEPVEKEEEPDPEPPPVTPQVGEGRKSRKFSRLSCLRRRRHPPKVGDDSDLSEGFESDSSHDSARASEGSDSGSDKSLEGGGTAFDAETDSEMNSQESRSDLEDMEEEEGTRSPTLEPPRGRSEAPDSLNGPLGPSEASIASNLQAMSTQMFQTKRCFRLAPTFSNLLLQPTTNPHTSASHRPCVNGDVDKPSEPASEEGSESEGSESSGRSCRNERSIQEKLQVLMAEGLLPAVKVFLDWLRTNPDLIIVCAQSSQSLWNRLSVLLNLLPAAGELQESGLALCPEVQDLLEGCELPDLPSSLLLPEDMALRNLPPLRAAHRRFNFDTDRPLLSTLEESVVRICCIRSFGHFIARLQGSILQFNPEVGIFVSIAQSEQESLLQQAQAQFRMAQEEARRNRLMRDMAQLRLQLEVSQLEGSLQQPKAQSAMSPYLVPDTQALCHHLPVIRQLATSGRFIVIIPRTVIDGLDLLKKEHPGARDGIRYLEAEFKKGNRYIRCQKEVGKSFERHKLKRQDADAWTLYKILDSCKQLT\LAQGAGEEDPSGMVTIITGLPLDNPSVLSGPMQAALQAAAHASVDIKNVLDFYKQWKEIG
3477	A	1	3902	MTEPRERRGYSVPPRPEVGTQATEWRVEESNFNKIFLKKDAELGRSNHLPTWDKPEDASWLPQSCLGGDAVATTGEIHEEKAWKTRALEVGQPAQRDIRRGELWGKEHGADQAIQETLEDLSSLERTLVVSESSPLGGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAGAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQVLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIVGPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSNDNSPAFAESSLALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPE\VLDTFSIDAKTGQVILRRPLDYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLDVNDNIPSIHVTWASQPSLVSEALPKDSFIALVMADDLDSGNNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSYRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDANDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTTIVARDADSGANGEPLYSIRSGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGSPPLQTRALLRVMFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLQAPFHLTPTLYRTLRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPEPQPATGQPRSRPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQILRSLVRLSVAAFAERNPVEELTVDSPPVQQISQLLSLLHQGQFQPKPNHRGNKYLAKPGGSRSAIPDTD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEE LSSLLDPSTGLALDRLSAPDPAWMARLSLPLTTN YRDNVISPDAAATEEPRTFQTFGKAEAPELSPTG TRLASTFVSEMSSLLEMLLEQRSSMPVEAASEAL RRLSVCGRTLSLDLATSAASGMKVQGDPGGKTG TEGKSRGSSSSSRCL
3478	A	13	1620	TLPPPGNSGCHRLCFPEFEFLQVTKMEFSGRKWR KLRLAGDQRNASYPHCLQFYLQPPSENISLIEFEN LAIDRVKLLKSVENLGVSYVKGTEQYQSKLESEL RKLKFSYRENLEDEYEPRRRDHISHFILRLAYCQS EELRRWFIQQEMDLLRFRFSILPKDKIQDFLKDSQ LQFEAISDEEKTLREQEIVASSPSLSGLKLGFESIY KIPFADALDLFRGRKVYLEDGFAYVPLKDIVAIIL NEFRAKLSKALALTARSLPAVQSDERLQPLLNHL SHSYTGQDYSTQGNVGKISLDQIDLLSTKSFPPC MRQLHKALRENHHLRHGGRMQYGLFLKGIGLT LEQALQFWKQEFIKGKMDPDKFDKGYSYNIRHS FGKEGKRTDYTPFSCLKIILSNPPSQGDYHGCPFR HSDPELLKQKLQSYKISPGGISQILDLVKGTHYQ V\ACQKYFEMIHTVDDCGFS\LSHPNQYFCESQRI LNGGKDIKKEPIQPETPQPKPSVQKTKDASSALA SLNSSLEMDMEGLEDYFSEDS
3479	A	698	138	RPELELWRLRSRSWRPLGVPRRCHRRNWKEPVR AQPLSVTVWAPRCQRP/QPPAPEPSSPNAAVPEAI PTPRAAASAALELPLGPAPVSVAPQAEAEARSTP GPAGSRLGPETFRQRFRQFRYQDAAGPREAFRQL REL/SPRQWLRPDI\RTKEQ\IVEMLVQEQLLAILP EAARARRIRRRTDVRITG
3480	A	117	2226	RRGSRSRGPFAEPAAPGGLCSSSEEKTEEGGMAV GLCKAMSQGLVTFRDVALDFSQEEWEWLKPSQ KDLYRDVMLENYRNLVWLGLSISKPNMISLLEQ GKEPWMVERKMSQGHCADWESWWEIEELSPK WFIDEDEISQEMVMERLASHGLECSSFREAWKY KGEFELHQGNAERHFMQVTAVKEISTGKRDNEF SN/IWEKHTPEISIFNTTES\PTIQQVHKFDIYDKLF PQNSVIIEYKRLHAEKESLIGNECEEFNQSTYLSK DIGIPPGEKPYESHDFSKLLSFHSLFTQHQTTHFG KLPHGYDECGDAFSCYSFFTQPQRIHSGEKPYAC NDCGKAFSHDFFLSEHQRTHIGEKPYECKECNKA FRQSAHLAQHQRIHTGEKPFACNECGKAFSRYAF LVEHQRIHTGEKPYECKECNKAFRQSAHLNQHQ RIHTGEKPYECNQCGKAFSRRIALTLHQRIHTGE KPFKCSECGKTFGYRSHLNQHQRIHTGEKPYECI KCGKFFRTDSQLNRHHRIHTGERPFECSKCGKAF SDALVLIHHKRSHAGEKPYECNKCGKAFSCGSY LNQHQRIHTGEKPYECSECGKAFHQILSLRLHQRI HAGEKPYKCNESQRVRRSELAVSRGLTTKPADT GPDSTLNAAKVAEPARAGTEAALRPALSVAESA TSLGPLHQGRRFPEAPAAHPGGTGFTVCAS
3481	A	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEAEG RLREKLFSGYDSSVRPAREVGDRVRVSVGLILAQ LISLNEKDEEMSTKVYLDLEWTDYRLSWDPAEH DGIDSLRITAESVWLPDVVLLNNNDGNFDVALDI SVVVSSDGSVRWQPPGIYRSSCSIQVTYFPFDWQ NCTMVFSSYSYDSSEVSLQTGLGPDGQGHQEIHI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HEGTFIENGQWENIHKPSRLIQPPGDPRGGREGQ RQEVIFYLIIRRKPLFYLVNVIAPCILITLLAIFVFY LPPDAGEKMGLSIFALLTLTVFLLLLADKVPETSL SVPIIIKYLMFTMVLVTFSVILSVVVLNLHHRSPH THQMPLWVRQIFIHKLPLYLRLKRPKPERDLMPE PPHCSSPGSGWGRGTDEYFIRKPPSDFLFPKPNRF QPELSAPDLRRFIDGPNRAVALLPELREVVSSISYI ARQLQEQEDHDALKEDWQFVAMVVDRLFLWTF IIFTSVGTL\VIFLDATYHLPPPDPFP
3482	A	1273	172	ERWDSGGADAEWYALADWTAVWLPRSDFYTR LQTGEGHVPALRLPAGMPPDSPRELVPKQAPCSP SDPALPWTLGHGNQPPAVVPEPQGPMGPAGVAA RPGRFFGVYLLYCLNPRYRVR\VYVGFTVNTARR VQQHNGGRKKGGA\GRTSGRGPWEMVLVVHGF PSSVAALRFEWAWQHPHASRRLAHVGPRLRGET AFAFHLRVLAHMLRAPPWARLPLTLRWVRPDLR QDLCLPPPPHVLLAFGPPPAQVPRPQRRRAGPFD DAEPEPDQGDPGACCSLCAQTIQDEEGPLCCPHP GCLLRAHVICLAEEFLQEEPGQLLPLEGQCPCCE KSLLWGDLIWLCQMDTEKEVEDSELEEAHWTD LLET
3483	A	230	3686	WRPWPCIDTSWNLQVAARTLRVSSAQCGLVPT MARVESPVPAARASLTGSCVLGQAMPLRGGAGP SPASHGPTHGPSDPRTCLPGRGAGGMRPHGRGA LGCCGLCSFYTCHGAAGDEIMHQDIVPLCAADIQ DQLKKRFAYLSGGRGQDGSPVITFPDYPAFSEIPD KEFQNVMTYLTSIPSLQDAGIGFILVIDRRRDKW TSVKASVLRIAASFPANLQLVLVLRPTGFFQRTLS DIAFKFNRDDFKMKVPVIMLSSVPDLHGYIDKSQ LTEDLGGTLDYCHSRWLCQRTAIESFALMVKQT AQMLQSFGTELAETELPNDVQST\SSVLCAHTEK KDKAKEDLRLALKEGHSVLESLRELQAEGSEPSV NQDQLDNQATVQRLLAQLNETEAAFDEFWAKH QQKLEQCLQLRHFEQGFREVKAILDAASQKIATF TDIGNSLAHVEHLLRDLANFQEKSGVFVERARA LSLTASSFIGNKHYAVDSIRPKCQELRHLCDQFSA EIARRRGLLSKSLELHRRLETSMKWCDEGIYLLA SQPVDKCQSQDGAEAALQEIEKFLETGAENKIQE LNAIYKEYESILNQDLMEHVRKVFQKQASMEEV FHRRQASLKKLAARQTRPVQPVAPRPEALAKSP CPSPGIRRGSENSSSEGGALRRGPYRRAKSEMSES RQGRGSAGEEEESLAILRRHVMSELLDTERAYVE ELLCVLEGYAAEMDNPLMAHLLSTGLHNKKDV LFGNMEEIYHFHNRIFLRELENYTDCPELVGRCF LERMEDFQIYEKYCQNKPRSESLWRQCSDCPFFQ ECQRKLDHKLSLDSYLLKPVQRITKYQLLLKEM LKYSRNCEGAEDLQEALSSILGILKAVNDSMHLI AITGYDGNLGDLGKLLMQGSFSVWTDHKRGHT KVKELARFKPMQRHLFLHEKAVLFCKKREENGE GYEKAPSYSYKQSLNMAAVGITENVKGDAKKFE IWYNAREEVYIVQAPTPEIKAAWVNEIRKVLTSQ LQACREASQHRALEQSQSLPLPAPTSTSPSRGNSR NIKKLEERKTDPLSLEGYVSSAPLTKPPEKGKGW SKTSHSLEAPEDDGGWSSAEEQINSSDAEEDGGL GPKKLVPGKYTVVADHEKGGPDALRVRSGDVV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ELVQEGDEGLW
3484	A	208	6103	VTMAQQAADKYLYVDKNFINNPLAQADWAAK KLVWVPSDKSGFEPASLKEEVGEEAIVELVENGK KVKVNKDDIQKMNPPKFSKVEDMAELTCLNEAS VLHNLKERYYSGLIYTYSGLFCVVINPYKNLPIYS EEIVEMYKGKKRHEMPPHIYAITDTAYRSMMQD REDQSILCTGESGAGKTENTKKVIQYLAYVASSH KSKKDQGELERQLLQANPILEAFGNAKTVKNDN SSRFGKFIRINFDVNGYIVGANIETYLLEKSRAIRQ AKEERTFHIFYYLLSGAGEHLKTDLLLEPYNKYR FLSNGHVTIPGQQDKDMFQETMEAMRIMGIPEEE QMGLLRVISGVLQLGNIVFKKERNTDQASMPDN TAAQKVSHLLGINVTDFTRGILTPRIKVGRDYVQ KAQTKEQADFAIEALAKATYERMFRWLVLRINK ALDKTKRQGASFIGILDIAGFEIFDLNSFEQLCINY TNEKLQQLFNHTMFILEQEEYQREGIEWNFIDFG LDLQPCIDLIEKPAGPPGILALLDEECWFPKATDK SFVEKVMQEQGTHPKFQKPKQLKDKADFCIIHY AGKVDYKADEWLMKNMDPLNDNIATLLHQSSD KFVSELWKDVDRIIGLDQVAGMSETALPGAFKT RKGMFRTVGQLYKEQLAKLMATLRNTNPNFVR CIIPNHEKKAGKLDPHLVLDQLRCNGVLEGIRICR QGFPNRVVFQEFRQRYEILTPNSIPKGFMDGKQA CVLMIKALELDSNLYRIGQSKVFFRAGVLAHLEE ERDLKITDVIIGFQACCRGYLARKAFAKRQQQLT AMKVLQRNCAAYLKLRNWQWWRLFTKVKPLL QVSRQEEEMMAKEEELVKVREKQLAAENRLTE METLQSQLMAEKLQLQEQLQAETELCAEAEELR ARLTAK\KQ\ELEEICHDLEARVEEEEERCQHLQA EKKKMQQNIQELEEQLEEEESARQKLQLEKVTT EAKLKKLEEEQIILEDQNCKLAKEKKLLEDRIAEF TTNLTEEEEKSKSLAKLKNKHEAMITDLEERLRR EEKQRQELEKTRRKLEGDSTDLSDQIAELQAQ\IA ELKMQLAKKEEELQAALARVEEEAAQKNMALK KIRELESQISELQEDLKCER\ASRNKAEKQKRDLG EELEALKTELEDTLDSTAAQQELRSKREQEVNIL KKTLEEEAKTHEAQIQEMRQKHSQAVEELAEQL EQTKRVKANLEKAKQTLENERGELANEVKVLLQ GKGDSEHKRKKVEAQLQELQVKFNEGERVRTEL ADKVTKLQVELDNVTGLLSQSDSKSSKLTKDFS ALESQLQDTQELLQEENRQKLSLSTKLKQVEDE KNS\FREQLEEEEEEAKHNLEKQIATLHAQVADM KKKMEDSVGCLETAEEVKRKLQKDLEGLSQRHE EKVAAYDKLEKTKTRLQQELDDLLVDLDHQRQ SACNLEKKQKKFDQLLAEEKTISAKYAEERDRA EAEAREKETKALSLARALEEAMEQKAELERLNK QFRTEMEDLMSSKDDVGKSVHELEKSKRAIEQQ VEEMKTQLEELEDELQATEDAKLRLEVNLQAM KAQFERDLQGRDEQSEEKKKQLVRQVREMEAE LEDERKQRSMAVAARKKLEMDLKDLEAHIDSA NKNRDEAIKQLRKLQAQMKDCMRELDDTRASR EEILAQAKENEKKLKSMEAEMIQLQEELAAAER AKRQAQQERDELADEIANSSGKGALALEEKRRL EARIAQLEEELEEEQGNTELINDRLKKANLQIDQI NTDLNLERSHAQKNENARQQLERQNKELKVKL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QEMEGTVKSKYKASITALEAKIAQLEEQLDNETKERQAACKQVRRTEKKLKDVLLQVDDERRNAEQYKDQADKASTRLKQLKRQLEEAEEEAQRANASRRKLQRELEDATETADAMNREVSSLKNKLRRGDLPFVVPRRMARKGAGDGSDEEVDGKADGAEAKPAE
3485	A	2	1782	CSTGVSKAPLTYLMSYGFELGWRKGNRAVACREDRGGESVGMGQESILSQVHWWEAPVEKTPGRDSEATIMSLRVHTLPTLLGAVVRPGCRELLCLLMITVTVGPGASGVCPTACICATDIVSCTNKNLSKVPGNLFRLIKRLDLSYNRIGLLDSEWIPVSFAKLNTLILRIINNITSISTGSFSTTPNLKCLDLSSNKLKT\VKNAVFQELKVLEVLLLYNNHISYLDPSAFGGLSQLQKLYLSGNFLTQFPMDLYVGRFKLAELMFLDVSYNRIPSMPMHHINLVPGKQLRGIYLHGNPFVCD\CSLVSLLVFWYRRHFSSVMDFKNDYTCRLWSDSRHSRQVLLLQDSFMNCSDSIINGSFRALGFIHEAQVGERLMVHCDSKTGNANTDFIWVGPDNRLLEPDKEMENFYVFHNGSLVIESPRFEDAGVYSCIAMNKQRLLNETVDVTINVSNFTVSRSHAHEAFNTAFTTLAACVASIVLVLLYLYLTPCPCKCKTKRQKNMLHQSNAHSSILSPGPASDASADERKAGAGKRVVFLEPLKDTAAGQNGKVRLFPSEAVIAEGILKSTRGKSDSDSVNSVFSDTPFVAST
3486	A	357	1173	GDPRETKVFPSRSFARNTVGVSHHQSHLFHTVSRIYVEDKHKILYCEVPKAGCSNWKRILMVLNGLASSAYNISHNAVHYGKHLKKLDSFDLKGIYTRLDTYTK\LVLVRDPMERLVSAFRDKFDHPNSYYHPVFGKAIIKKYRPNACEEALINGSGVKFKEFIHYLLDSHRPVGMDIHWEKVSKLCYPCLINYDFVGKFETLEEDANYFLQMIGAPKELKFPNFKDRHSSDERTNAQVVRQYLKDLTRTERQLIYDFYYLDYLMFNYTTPFL
3487	A	2	3281	CDKSGAVPFSTTRSPRRPSPRSAGPSLSSVSPRSQLWASSGLSEEHAAPLLPAWPRHPCPPSLTPGPSMAQGAMRFCSEGDCAISPPRCPRRWLPEGPVPQSPPASMYGSTGSLLRRVAGPGPRGRELGRVTAPCTPLRGPPSPRVAPSPWAPSSPTGQPPPGAQSSVVIFRFVEKASVRPLNGLPAPGGLSRSWDLGGVSPPRPTPALGPGSNRKLRLEASTSDPLPARGGSALPGSRNLVHGPPAPPQVGADGLYSSLPNGLGDPPERLATLFGGPADTGFLNQGDTWSSPREVSSHAQRIARAKWEFFYGSLDPPSSGAKPPEQAPPSPPGVGSRQGSGVAVGRAAKYSETDLDTVPLRCYRETDIDEVLAEREEADSAIESQPSSEGPPGTAYPPAPRPGPLPGPHPSLGSGNEDEDDDEAGGEEDVDDEVFEASEGARPGSRMPLKSPVPFLPGTSPSADGPDSFSCVFEAILESHRAKGTSYTSLASLEALASPGPTQSPFFTFELPPQPPAPRPDPPAPAPLAPLEPDSGTSSAADGPWTQRGEEEEAEARAKLAPGREPPSPCHSEDSLGLGAAPLGSEPPLSQLVSDSDSELDSTERLALGSTDTLSNGQKADLEAAQRLAKRLYRLDGFRKADVARHLGKNNDFSKLVAGEYLKFFVFTGMTLDQALRVFLKELALMGETQERERVLAHFSQRYFQCNPEALSSEDGAHTLTCALMLLNTDLHGHNIGKRMTCGDFIG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NLEGLNDGGDFPRELLKALYSSIKNEKLQWAIDE EELRRFLSELADPNPKVIKRISGGSGSGSSPFLDLT PEPGAAVYKHGALVRKVHADPDCRKTPRGKRG WKSFHGILKGMILYLQKEEYKPGKALSETELKN AISIHHALATRAS\NYSKRPHVFYLRTADWRVFL FQAPSLEQMQSWITRINVVAAMFSAPPFPAAVSS QKKFSRPLLPSAATRLSQEEQVRTHEAKLKAMA SELREHRAAQLGKKGRGKEAEEQRQKEAYLEFE KSRYSTYAALLRVKLKAGSEELDAVEAALAQAG STEDGLPPSHSSPSLQPKPSSQPRAQRHSSEPRPG AGSGRRKP
3488	A	441	1968	GTETPHCWGRGTAGLRRELDREERDGPGTATMS FPHFGHPYRGAFQFL\ASASSSTTCCESTLRSVSY VASGSTPAPALCCAP\YDSRLLGSARPELGAALGI YGAPYAAAAAAQSYPGYLPYSPEPPSLYGALNP QYEFKEAAGSFTSSLAQPGAYYPYERTLGQYQY ERYGAVELSGAGRRKNATRETTSTLKAWLNEHR KNPYPTKGEKIMLAIITKMTLTQVSTWFANARRR LKKENKMTWAPKNKGGEERKAEGGEEDSLGCL TADTKEVTASQEARGLRLSDLEDLEEEEEEEEA EDEEVVATAGDRLTEFRKGAQSLPGPCAAAREG RLERRECGLAAPRFSFNDPSGSEEADFLSAETGSP RLTMHYPCLEKPRIWSLAHTATASAVEGAPPARP RPRSPECRMIPGQPPASARRLSVPRDSACDESSCI PKAFGNPKFALQGLPLNCAPCPRRSEPVVQCQYP SGAEGSGPPAALGVSMQKTPTYRPARQLHTLCH SSLP
3489	A	718	2073	IAAYHKALSYRGHVHANNRGTNNVHFTPPPSPS RGILPMNPRNMMNHSQVGQGIGIPSRTNSMSSSG LGSPNRSSPSIICMPKQQPSRQPFTVNSMSGFGMN RNQAFGMNNSLSSNIFNGTDGSENVTGLDLSDFP ALADRNRREGSGNPTPLINPLAGRAPYVGMVTK PANEQSQDFSIHNEDFPALPGSSYKDPTSSNDDSK SNLNTSGKTTSSTDGPKFPGDKSSTTQNNNQQKK GIQVLPDGRVTNIPQGMVTDQFGMIGLLTFIRAA ETDPGMVHLALGSDLTTLGLNLNSPENLYPKFAS PWASSPCRPQDIDFHVPSEYLTNIHIRDKLFFFS W/TAIKLGRYGEDLLFYLYYMNGGDVLQLLAAV ELFNRDWRYHKEERVWITRAPGMEPTMKTNTY ERGTYYFFDCLNWRKVAKEFHLEYDKLEERPHL PSTFNYNPAQQAF
3490	A	2	2833	FVAKMATSQYFDFAQGGGPQYSTQAPTLPLPTV GASYTGQPTPGMDPAVNPAFPPAAPAGYGGYQP HSGQDFAYGSRPQEPVPTATTMATYQDSYSYGQ SAAARSYEDRPYFQSAALQSGRMTAADSGQPGT QEACGQPSPHGSHSHAQPPQQAPIVESGQPASTL SSGYTYPTATGVQPESSASIVTSYPPPSYNPTCTA YTAPSYPNYDASVYSAASPFYPPAQPPPPPGPPQ QLPPPPAPAGSGSSPRADSKPPLPSKLPRPKAGPR QLQLHYCDICKISCAGPQTYREHLGGQKHRKKE AAQKTGVQPNGSPRGVQAQLHCDLCAVSCTGA DAYAAHIRGSKHQKVFKLHAKLGKPIPTLEPALA TESPPGAEAKPTSPTGPSVCASSRPALAKRPVASK ALCEGPPEPQAAGCRPQWGKPAQPKLEGPGAPT QGGSKEAPAGCSDAQPVGPEYVEEVFSDEGRVL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RFHCKLCECSFNDLNAKDLHVRGRRHRLQYRKKVNPDLPIATEPSSRARKVLEERMRKQRHLAEERLEQLRRWHAERRRLEEEPPQDVPPHAPPDWAQPLLMGRPESPASAPLQPGRRPASSDDRHVMCKHATIYPTEQELLAVQRAVSHAERALKLVSDTLAEEDRGRREEEGDKRSSVAPQTRVLKGVMRVGILAKGLLLRGDRNVRLALLCSEKPTHSLLRRIAQQLPRQLQMVTEDEYEVSSDPEANIVISSCEEPRMQVTISVTSPLMREDPSTDPGVEEPQADAGDVLSPKKCLESLAALRHARWFQARASGLQPCVIVIRVLRDLCRRVPT\WGALPAWAMELLVEKAVSSAAGPLGPGDAVRRVLECVATGTLLTDGPGLQDPCERDQTDALEPMTLQEREDVTASAQHALRMLAFRQTHKVLGMDLLPPRHRLGARFRKRQRGPGEGEEGAGEKKRGRRGGEGLV
3491	A	2	1321	FVGDGALSGCRRGRAPRVPSMAGSLPPCVVDCGTGYTKLGYAGNTEPQFIIPSCIAIRESAKVVDQAQRRVLRGVDDLDFFIGDEAIDKPTYATKWPIRHGIIEDWDLMERFMEQVVFKYLRAEPEDHYFLMTEPPLNTPENREYLAEIMFESFNVPGLYIAVQAVLALAASWTSRQVGERTLTGIVIDSGDGVTHVIPVAEGYVIGSCIKHIPIAGRDITYFIQQLLREREVGIPPEQSLETAKAIKEKYCYICPDIVKEFAKYDVDPRKWIKQYTGINAINQKKFVIDVGYERFLGPEIFFHPEFANPDFMESISDVVDEVIQNCPIDVRRPLYKNVVLSGGSTMFRDFGRRLQRDLKRVVDARLRLSEELSGG\RIKPKPVEVQVVTHHMQRYAV\WFGG\SMLASTPEFFQVCHTKKDYEEYGPSICRHNPVFGVMS
3492	A	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSRGSREESPAPSRAPASASLWRRLVVVEAKMAAHAAAAAQAAAAQAAHAEAADSWYLALLGFAEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEARTHLQLGSVLYHHTKNSEQARSHLEKAWLISQQIPQFEDVKFEAASLLSELYCQENSVDAAKPLLRKAIQISQQTPYWHCRLLFQLAQLHTLEKDLVSACDLLGVGAEYARVVGSEYTRALFLLSKGMLLLMERKLQEVHPLLTLCGQIVENWQGNPIQKESLRVFFLVLQVTHYLDAGQVKSVKPCLKQLQQCIQTISTLHDDEILPSNPADLFHWLPKEHMCVLVYLVTVMHSMQAGYLEKAQKYTDKALMQLEKLKMLDCSPILSSFQVILLEHIIMCRLVTGHKATALQEISQVCQLCQQSPRLFSNHAAQLHTLLGLYCVSVNCMDNAEAQFTTALRLTNHQELWAFIVTNLASVYIREGNRHQEVV\LYSLLERINPDHSFPVSSHCLRAAAFYVRGLFSFFQGRYNEAKRFLRETLKMSNAEDLNRLTACSLVLLGHIFYVLGNHRESNNMVVPAMQLASKIPDMSVQLWSSALLRDLNKACGNAMDAHEAAQMHQNFSQQLLQDHIEACSLPEHNLITWTDGPPPVQFQAQNGPNTSLASLL
3493	A	3	2024	PNGVALLHLPGAAVIPNTNYMFQDALGGRSRGSREESPAPSRAPASASLWRRLVVVEAKMAAHAAAAAQAAAAQAAHAEAADSWYLALLGFAEHFRTSSPPKIRLCVHCLQAVFPFKPPQRIEARTHLQLGSVLYHHTKNSEQARSHLEKAWLISQQIPQFEDVKFEAASLLSELYCQENSVDAAKPLLRKAIQISQQTPY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				WHCRLLFQLAQLHTLEKDLVSACDLLGVGAEY ARVVGSEYTRALFLLSKGMLLLMERKLQEVHPL LTLCGQIVENWQGNPIQKESLRVFFLVLQVTHYL DAGQVKSVKPCLKQLQQCIQTISTLHDDEILPSNP ADLFHWLPKEHMCVLVYLVTVMHSMQAGYLE KAQKYTDKALMQLEKLKMLDCSPILSSFQVILLE HIIMCRLVTGHKATALQEISQVCQLCQQSPRLFS NHAAQLHTLLGLYCVSVNCMDNAEAQFTTALR LTNHQELWAFIVTNLASVYIREGNRHQEVV\LYS LLERINPDHSFPVSSHCLRAAAFYVRGLFSFFQGR YNEAKRFLRETLKMSNAEDLNRLTACSLVLLGHI FYVLGNHRESNNMVVPAMQLASKIPDMSVQLW SSALLRDLNKACGNAMDAHEAAQMHQNFSQQL LQDHIEACSLPEHNLITWTDGPPPVQFQAQNGPN TSLASLL
3494	A	2	1615	VLRGQRGPAGGLAEERRRGRNEWRIHDVTTAPF PGLVQRRSRLLIVSQVRYFLKNKVSPDLCNEDGL TALHQCCIDNFEEIVKLLLSHGANVNAKDNELW TPLHAAATCGHINLVKILVQYGADLLAVNSDGN MPYDLCEDEPTLDVIETCMAYQGITQEKINEMRV APEQQMIADIHCMIAAGQDLDWIDAQGATLLHI AGANGYLRAAELLLDHGVRVDVKDWDGWEPL HAAAFWGQMQMAELLVSHGAN\LNARTSMDE MPIDLCEEEEFKVLLLELK\HKHDVIMKSQLRHK SSLSRRTSHRQAS/SVGKVVRRTQPVGTGPNL\YR KEYE/GEEAILWQRSA\AEDQRTSTYNGDIRET\R TDQENKDPNPRLEK\PVLLSEFPTKIPRGELDMPV ENGLRAPVSAYQYALANGDVWKVHEVPDYSM AYGNPGVADATPPWSSYKEQSPQTLLELKRQRA AAKLLSHPFLSTHLGSSMARTGESSSEGKAPLIG GRTSPYSSNGTSVYYTVTSGDPPLLKFKAPIEEM EEKVHGCCRIS
3495	A	327	1078	APMADTTPNGPQGAGAVQFMMTNKLDTAMWL SRLFTVYCSALFVLPLLGLHEAASFYQRALLANA LTSALRLHQRLPHFQLSRAFLAQALLEDSCHYLL YSLIFVNSYPVTMSIFPVLLFSLLHAATYTKKVL\ DARG\SNSLPLLR\SVLDKLSANQQNILKFIACNEI FLMPATVFMLFSGQGSLLQPFIYYRFLTLRYSSRR NPYCRTLFNELRIVVEHIIMKPACPLFVRRLCLQS IAFISRLAPTVP
3496	A	3	2867	SSRTREMEEKEILRRQIRLLQGLIDDYKTLHGNAP APGTPAASGWQPPTYHSGRAFSARYPRPSRRGYS SHHGPSWRKKYSLVNRPPGPSDPPADHAVRPLH GARGGQPPVPQQHVLERQVQLSQGQNVVIKVKP PSKSGSASASGAQRGSLEEFEDTPWSDQRPREGE GEPPRGQLQPSRPTRARGTCSVEDPLLVCQKEPG KPRMVKSVGSVGDSPREPRRTVSESVIAVKASFP SSALPPRTGVALGRKLGSHSVASCAPQLLGDRRV DAGHTDQPVPSGSVGGPARPASGPRQAREASLV VTCRTNKFRKNNYKWVAASSKSPRVARRALSPR VAAENVCKASAGMANKVEKPQLIADPEPKPRKP ATSSKPGSAPSKYKWKASSPSASSSSSFRWQSEA GSKDHASQLSPVLSRSPSGD\RPALAHSGLKPLSG ETPLSAYKVKTRTKIIRRRGSTSLPGDKKSGTSPA ATAKSHLSLRRRQALRGKSSPVLKKTPNKGLVQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VTKHRLCRLPPSRAHLPTKEASSLHAVRTAPTSK VIKTRYRIVKKTPASPLSAPPFPLSLPSWRARRLS LSRSLVLNRLRPVASGGGKAQPGSPWWRSKGYR CIGGVLYKVSANKLSKTSGQPSDAGSRPLLRTGR LDPAGSCSRSLASRAVQRSLAIIRQARQRREKRK EYCMYYNRFGRCNRGERCPYIHDPEKVAVCTRF VRGTCKKTDGTCPFSHHVSKEKMPVCSYFLKGI CSNSNCPYSHVYVSRKAEVCSDFLKGYCPLGAK CKKKHTLLCPDFARRGACPRGAQCQLLHRTQKR HSRRAATSPAPGPSDATARSRVSASHGPRKPSAS QRPTRQTPSSAALTAAAVAAPPHCPGGSASPSSS KASSSSSSSSPPASLDHE\APSLQEAALAAACSN RLCKLPSFISLQSSPSPGAQPRVRAPRAPLTKDSG KPLHIKPRL
3497	A	1586	141	ATARDLGCARRIDRVVMESTPSRGLNRVHLQCR NLQEFLGGLSPGVLDRLYGHPATCLAVFRELPSL AKNWVMRMLFLEQPLPQAAVALWVKKEFSKA QEESTGLLSGLRIWHTQLLPGGLQGLILNPIFRQN LRIALLGGGKAWSDDTSQLGPDKHARDVPSLDK YAEERWEVVLHFMVGSPSAAVSQDLAQLLSQA GLMKSTEPGEPPCITSAGFQFLLLDTPAQLWYFM LQYLQTAQSRGMDLVEILSFLFQLSFSTLGKDYS VEGMSDSLLNFLQHLREFGLVFQRKRKSRRYYP T/RALAINLSSGVSGAGGTVHQPGFIV\VETNYRL YAYTESELQIALIALFSEMLYPFP\NMVV\ARVTR\ ESVQQAIASGITAQQIIHFLRTRAHPVMLKQTPVL PPTITDQIRLWELERDRLRFTEGVLYNQFLSQVDF ELL\LAHAPKLGVLVFE/NTPAKRLMVVTPAGHS DVKRFWKRQKHSS
3498	A	790	190	RDLGPAALMTASASSFSSSQGVQQPSIYSFSQITR SLFLSNGVAANDKLLLSSNRITAIVNASVGSGQRI LRG\LQYIKVPVTDARDSRLYDFFDPIADLIHTVS MRQGRTLLNCMAG\MSRSASLCLAYLMKYHSM S\LLDAHTWA/TKSRRPIIRPNNGFWEQLINYEFK LFNNNTVRMINSPVGNIPDIYEKDLRMMISM
3499	A	31	1586	TAGFLLAPLEMQRLLTPVKRILQLTRAVQETSLT PARLLPVAHQRFSTASAVPLAKTDTWPKDVGIL ALEVYFPAQYVDQTDLEKYNNVEAGKYTVGLG QTRMGFCSVQEDINSLCLTVVQRLMERIQLPWD SVGRLEVGTETIIDKSKAVKTVLMELFQDSGNTD IEGIDTTNACYGGTASLFNAANWMESSSWDGRY AMVVCGDIAVYPSGNARPTGGAGAVAMLIGPK APLALERGLRGTHMENVYDFYKPNLASEYPIVD GKLSIQCYLRALDRCYTSYRKKIQNQWKQAGSD RPFTLDDLQYMIFHTPFCKMVQKSLARLMFNDF LSASSDTQTSLYKGLEAFGGLKLEDTYTNKDLD KALLKASQDMFDKKTKASLYLSTHNGNMYTSSL YGCLASLLSHHSAQELAGSRIGAFSYGSGLAASF FSFRVSQDAAPGSPL\DKLVSSTSDLPKRLASRKC VSPEEFTEIMNQREQFYHKVNFSPPGDTNSLFPGT WYLERVDEQHRRKYARRPV
3500	A	185	2692	MLPTEVPQSHPGPSALLLLQLLLPPTSAFFPNIWS LLAAPGSITHQDLTEEAALNVTLQLFLEQPPPGRP PLRLEDFLGRTLLADDLFAAYFGPGSSRRFRAAL GEVSRANAAQDFLPTSRNDPDLHFDAERLGQGR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ARLVGALRETVVAARALDHTLARQRLGAALHA LQDFYSHSNWVELGEQQPHPHLLWPRQELQNLA QVADPTCSDCEELSCPRNWLGFTLLTSGYFGTHP PKPPGKCSHGGHFDRSSSQPPRGGINKDSTSPGFS PHHMLHLQAAKLALLASIQAFSLLRSRLGDRDFS RLLDITPASSLSFVLDTTGSMGEEINAAKIQARHL VEQRRGSPMEPVHYVLVPFHDPGFGPVFTTSDPD SFWQQLNEIHALGGGDEPEMCLSALQLALLHTPP LSDIFVFTDASPKDAFLTNQVESLTQERRCRVTFL VTEDTSRVQGRARREILSPLRFEPYKAVALASGG EVIFTKDQHIRDVAAIVGESMAALVTLPLDPPVV VPGQPLVFSVDGLLQKITVRIHGDISSFWIKNPAG VSQGQEEGGGPLGHTRRFGQFWMVTMDDPPQT GTWEIQVTAEDTPGVRVQAQTSLDFLFHFGIPME DGPHPGLYPLTQPVAGLQTQLLVEVTGLGSRAN PGDPQPHFSHVILRGVPEGAELGQVPLEPVGPPE RGLLAASLSPTLLSTPRPFSLELIGQDAAGRRLHR AAPQPSTVVPVLLELSGPSGFLAPGSKVPLSLRIA SFSGPQDLDLRTFVNPSFSLTSNLSRAHLELNESA WGRLWLEVPDSAAPDSVVMVTVTAGGREANPV PPTHAFLRLLVSAPAPQDRH
3501	A	1245	5815	RRAHPSHSRLSPYLSVSRDPYFFVTVSRTILTLSA PAPPRRTPAPSMGTALLQRGGCFLLCLSLLLLGC WAELGSGLEFPGAEGQWTRFPKWNACCESEMSF QLKTRSARGLVLYFDDEGFCDFLELILTRGGRLQ LSFSIFCAEPATLLADTPVNDGAWHSVRIRRQFR NTTLFIDQVEAKWVEVKSKRRDMTVFSGLFVGG LPPELRAAALKLTLASVREREPFKGWIRDVRVNS SQVLPVDSGEVKLDDEPPNSGGG\SPCEAGEEGE GGVCLNGGVCSVVDDQAVCDCSRTGFRGKDCS QEDNNVEGLAHLMMGDQGKEEYIATFKGSEYF CYDLSQNPIQSSSDEITLSFKTLQRNGLMLHTGKS ADYVNLALKNGAVSLVINLGSGAFEALVEPVNG KFNDNAWHDVKVTRNLRQHSGIGHAMVTISVD GILTTTGYTQEDYTMLGSDDFFYVGGSPSTADLP GSPVSNNFMGCLKEVVYKNNDVRLELSRLAKQ GDPKMKIHGVVAFKCENVATLDPITFETPESFISL PKWNAKKTGSISFDFRTTEPNGLILFSHGKPRHQ KDAKHPQMIKVDFFAIEMLDGHLYLLLDMGSGT IKIKALLKKVNDGEWYHVDFQRDGRSGTISVNT LRTPYTAPGESEILDLDDELYLGGLPENKAGLVF PTEVWTALLNYGYVGCIRDLFIDGQSKDIRQMA EVQSTAGVKPSCSKETAKPCLSNPCKNNGMCRD GWNRYVCDCSGTGYLGRSCEREATVLSYDGSM FMKIQLPVVMHTEAEDVSLRFRSQRAYGILMAT TSRDSADTLRLELDAGRVKLTVNLDCIRINCNSS KGPETLFAGYNLNDNEWHTVRVVRRGKSLKLT VDDQQAMTGQMAGDHTRLEFHNIETGIITERRY LSSVPSNFIGHLQSLTFNGMAYIDLCKNGDIDYC ELNARFGFRNIIADPVTFKTKSSYVALATLQAYT SMHLFFQFKTTSLDGLILYNSGDGNDFIVVELVK GYLHYVFDLGNGANLIKGSSNKPLNDNQWHNV MISRDTSNLHTVKIDTKITTQITAGARNLDLKSDL YIGGVAKETYKSLPKLVHAKEGFQGCLASVDLN G\RLP\DLISDGSFSCNGTDSRRGMWKGPSTT\CQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EDSCSNQGVCLQQWDGFSCDCSMTSFSGPLCND PGTTYIFSKGGGQITYKWPPNDRPSTRADRLAIGF STVQKEAVLVRVDSSSGLGDYLELHIHQGKIGVK FNVGTDDIAIEESNAIINDGKYHVVRFTRSGGNA TLQVDSWPVIERYPAGRQLTIFNSQATIIIGGKEQ GQPFQGQLSGLYYNGLKVLNMAAENDANIAIVG NVRLVGEVPSSMTTESTATAMQSEMSTSIMETTT TLATSTARRGKPPTKEPISQTTDDILVASAECPSD DEDIDPCEPSSGGLANPTRAGGREPYPGSAEVIRE SSSTTGMVVGIVAAAALCILILLYAMYKYRNRDE GSYHVDESRNYISNSAQSNGAVVKEKQPSSAKSS NKNKKNKDKEYYV
3502	A	394	72	KPAHLPFTVIIMPKRKPSEGAMSDKVKA/KFELQ RRSAGLFSKPTPPKPETRPKKDPANQRQKLPKVR KGKADA/SKEGNSPAEERCSMVQTQKVEGWRSG SELPVALSF
3503	A	43	3358	SGGRGPVRVRSEQLSPSAEQVSQISQISLGRRPLS SLPPPPSRALAPTRAPDTALTIMEVAEVESPLNPS CKIMTFRPSMEEFREFNKYLAYMESKGAHRAGL AKVIPPKEWKPRQCYDDIDNLLIPAPIQQMVTGQ SGLFTQYNIQKKAMTVKEFRQLANSGKYCTPRY LDYEDLERKYWKNLTFVAPIYGADINGSIYDEGV DEWNIARLNTVLDVVEEECGISIEGVNTPYLYFG MWKTTFAWHTEDMDLYSINYLHFGEPKSWYAIP PEHGKRLERLAQGFFPSSSQGCDAFLRHKMTLIS PSVLKKYGIPFDKITQEAGEFMITFPYGYHAGFN HGFNCAESTNFATVRWIDYGKVAKLCTCRKDM VKISMDIFVRKFQPDRYQLWKQGKDIYTIDHTKP TPASTPEVKAWLQRRRKVRKASRSFQCARSTSK RPKADEEEEVSDEVDGAEVPNPDSVTDDLKVSE KSEAAVKLRNTEASSEEESSASRMQVEQNLSDHI KLSGNSCLSTSVTEDIKTEDDKAYAYRSVPSISSE ADDSIPLSTGYEKPEKSDPSELSWPKSPESCSSVA ESNGVLTEGEESDVESHGNGLEPGEIPAVPSGER NSFKVPSIAEGENKTSKSWRHPLSRPPARSPMTL VKQQAPSDEELPEVLSIEEEVEETESWAKPLIHL WQTKPPNFAAEQEYNATVARMKPHCAICTLLMP YHKPDSSNEENDARWETKLDEVVTSEGKTKPLIP EMCFIYSEENIEYSPPNAFLEEDGTSLLISCAKCC VRVHASCYGIPSHEICDGWLCARCKRNAWTAEC CLCNLRGGALKQTKNNKWAHVMCAVAVPEVR FTNVPERTQIDVGRIPLQRLKLKCIFCRHRVKRVS GACIQCSYGRCPASFHVTCAHAAGVL\MEPDDW PYVVNITCFRHKVNPNVKSKACEKVISVGQTVIT KHRNTRYYSCRVMAVTSQTFYEVMFDDGSFSRD TFPEDIVSRDCLKLGPPAEGEVVQVKWPDGKLY GAKYFGSNIAHMYQVEFEDGSQIAMKREDIYTL DEELPKRVKARFVSAGRCHLGTCQVNSLSSPHVS QAQQETYLGFWINSKKSQCNIFLSGTY
3504	A	1124	139	RGEEQFDAEFRRFACLGFGERLQEFSRLLRAVHR SRAWTCYLAIRMLMATCCPSPTTTACTGPWQRA PPLRLLVQKREADSSGLAFASNSLQRRKKGLLLR PVAPLRTRPPLLISLPQDFRQVSSVIDVDLLPETH RRVRLHKHGSDRPLGFYIRDGMSVRVAPQG\LER VPGIFISRLVRGGLAESTGLLAVSDEILEVNGIEV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AGKTLNQVTDMMVANSHN\LIVTVKPANQRNN VVRGASGRLTGPPSAGPGPAEPDSDDDSSDLVIE NRQPPSSNGLSQGPPCWDLHPGCRHPGTRSSLPS LDDQEQASSGWGSRIRGDGSGFSL
3505	A	3	2898	SCRSATSQSGCGGGRSWLCSSLKMAAQPPRGIRL SALCPKFLHTNSTSHTWPFSAVAELIDNAYDPDV NAKQIWIDKTVINDHICLTFTDNGNGMTSDKLH KMLSFGFSDKVTMNGHVPVGLYGNGFKSGSM\R LGKDAIVFTKNGESMSVGLLSQTYL\EVIKAEHV VVPIVAFNKHRQMINLAESKASLAAILEHSLFSTE QKLLAELDAIIGKKGTRIIIWNLRSYKNATEFDFE KDKYDIRIPEDLDEITGKKGYKKQERMDQIAPES DYSLRAYCSILYLKPRMQIILRGQKVKTQLVSKS LAYIERDVYRPKFLSKTVRITFGFNCRNKDHYGI MMYHRNRLIKAYEKVGCQLRANNMGVGVVGII ECNFLKPTHNKQDFDYTNEYRLTITALGEKLND YWNEMKVKKNTEYPLNLPVEDIQKRPDQTWVQ CDACLKWRKLPDGMDQLPEKWYCSNNP\DPQFR NCEVPEEPEDEDLVHPTYEKTYKKTNKEKFRIRQ PEMIPRINAELLFRPT\ALSTPS\FSSPKESVSKR/RH LSEGTNSYATRLLNNHQVPPQSEPESNSLKRRLS TRSSILNAKNRRL\SSQF\ENSVYKG\DDDDEDVII LEENSTPKPAVDHDIDMKSEQSHVEQGGVQVEF VGDSEPCGQTGSTSTSSSRCDQGNTAATQTEVPS LVVKKEETVEDEIDVRNDAVILPSCVEAEAKIHE TQETTDKSADDAGCQLQELRNQLLLVTEEKENY KRQCHMFTDQIKVLQQRILEMNDKYVKKETCH QSTETDAVFLLESINGKSESPDHMVSQYQQALEE IERLKKQCSALQHVKAECSQCSNNESKSEMDEM AVQLDDVFRQLDKCSIERDQYKSEVELLEMEKS QIRSQCEELKTEVEQLKSTNQQTATDVSTSSNIEE SVNHMDGESLKLRSLRVNVGQLLAMIVPDLDLQ QVNYDVDVVDEILGQVVEQMSEISST
3506	A	2	2120	RPPEAGGRYRAGGRRQAAKPSRPPLPSRRRLPQG GRTRRAMDRPAAAAAAGCEGGGGPNPGPAGGR RPPRAAGGATAGSRQPSVETLDSPTGSHVEWCK QLIAATISSQISGSVTSENVSRDYKALRDGNKLA QMEEAPLFPGESIKAIVKDVMYICPFMGAVSGTL TVTDFKLYFKNVERDPHFILDVPLGVISRVEKIGA QSHGDNSCGIEIVCKDMRNLRLAYK\QEEQSKLG IFENLNKHAFPLSNGQALFAFSYKEKFPINGWKV YDPVSEYKRQGLPNESWKISKINSNYEFCDTYPA IIVVPTSVKDDDLSKVAVFLAKGRVPVLSWIHPE SQATITRCSQPLVGPNDKRCKEDEKYLQTIMDAN AQSHKLIIFDARQNSVADTNKTKGGGYESESAYP NAELVFLEIHNIHVMRESLRKLKEIVYPSIDEARW LSNVDGTHWLEYIRMLLAGAVRIADKIESGKTSV VVHCSDGWDRTAQLTSLAMLMLDSYYRTIKGFE TLVEKEWISFGHRFALRVGHGNDNHADADRSPIF LQFVDCVWQMTRQFPSAFEFNELFLITILDHLYS CLFGTFLCNCEQQRFKEDVYTKTISLWSYINSQL DEFSNPFFVNYENHVLYPVASLSHLELWVNYYV RWNPRMRPQMPIHQNLKELLAVRAELQKRVEG LQREVATRAVSSSSERGSSPSHFATSVHTLV
3507	A	1	2169	GSSIKIRLTVLCAKNLAKKDFFRLPDPF\AKIVVD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GSGQCHSTDTVKNTLDPKWNQHYDLYVGKTDSI TISVWNHKKIHKKQGAGFLGCVRLLSNAISRLKD TGYQRLDLCKLNPSDTDAVRGQIVVSLQTRDRIG TGGSVVDCRGLLENEGTVYEDSGPGRPLSCFME EPAPYTDSTGAAAGGGNCRFVESPSQDQRLQAQ RLRNPDVRGSLQTPQNRPHGHQSPELPEGYEQRT TVQGQVYFLHTQTGVSTWHDPRIPRDLNSVNCD ELGPLPPGWEVRSTVSGRIYFVDHNNRTTQFTDP RLHHIMNHQCQLKEPSQPLPLPSEGSLEDEELPA QRYERDLVQKLKVLRHELSLQQPQAGHCRIEVS REEIFEESYRQIMKMRPKDLKKRLMVKFRGEEG LDYGGVAREWLYLLCHEMLNPYYGLFQYSTDNI YMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGH YINGGFTVPFYKQLLGKPIQLSDLESVDPELHKSL VWILENDITPVLDHTFCVEHNAFGRILQHELKPN G\RNVPVTEENKKEYVRLYVNWRFMRGIEAQFL ALQKGFNELIPQHLLKPFDQKELELIIGGLDKIDL NDWKSNTRLKHCVADSNIVRWFWQAVETFDEE RRARLLQFVTGSTRVPLQGFKALQGSTG\AAGPR LFTIHLIDANTDNLRKAHTCFNRIDIPPYESYEKL YEKLLTAVEETCGFAVE
3508	A	3	6388	ILYINPADLGWNPPVSSWIEKREIQTERANLTILF DKYLPTCLDTLRTRFKKIIPIPEQSMVQMVCHLLE CLLTTEDIPADCPKEIYEHYFVFAAIWAFGGAMV QDQLVDYRAEFSKWWLTEFKTVKFPSQGTIFDY YIDPETKKFEPWSKLVPQFEFDPEMPLQACLVHT SETIRVCYFMERLMARQRPVMLVGTAGTGKSVL VGAKLASLDPEAYLVKNVPFNYYTTSAMLQAVL EKPLEKKAGRNYGPPGNKKLIYFIDDMNMPEVD AYGTVQPHTIIRQHLDYGHWYDRSKLSLKEITNV QYVSCMNPTAGSFTINPRLQRHFSVFVLSFPGAD ALSSIYSIILTQHLKLGNFPASLQKSIPPLIDLALAF HQKIATTFLPTGIKFHYIFNLRDFANIFQGILFSSV ECVKSTWDLIRLYLHESNRVYRDKMVEEKDFDL FDKIQTEVLKKTFDDIEDPVEQTQSPNLYCHFAN GIGEPKYMPVQSWELLTQTLVEALENHNEVNTV MDLVLFEDAMRHVCHINRILESPRGNALLVGVG GSGKQSLTRLAAFISSMDVFQITLRKGYQIQDFK MDLASLCLKAGVKNLNTVFLMTDAQVADERFL VLINDLLASGEIPDLYSDDEVENIISNVRNEVKSQ GLVDNRENCWKFFIDRIRRQLKVTLCFSPVGNKL RVRSRKFPAIVNCTAIHWFHEWPQQALESVSLRF LQNTEGIEPTVKQSISKFMAFVHTSVNQTSQSYLS NEQRYNYTTPKSFLEFIRLYQSLLHRHRKELKCK TERLENGLLKLHSTSAQVDDLKAKLAAQEVELK QKNEDADKLIQVVGVETDKVSREKAMADEEEQ KVAVIMLEVKQKQKDCEEDLAKAEPALTAAQA ALNTLNKTNLTELKSFGSPPLAVSNVSAAVMVL MAPRGRVPKDRSWKAAKVTMAKVDGFLDSLIN FNKENIHENCLKAIRPYLQDPEFNPEFVATKSYA AAGLCSWVINIVRFYEVFCDVEPKRQALNKATA DLTAAQEKLAAIKAKIAHLNENLAKLTARFEKA TADKLKCQQEAEVTAVTISLANRLVGGLASENV RWADAVQNFKQQERTLCGDILLITAFISYLGFFT KKYRQSLLDRTWRPYLSQLKTPIPVTPALDPLRM

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LMDDADVAAWQNEGLPADRMSVENATILINCE RWPLMVDPQLQGIKWIKNKYGEDLRVTQIGQKG YLQIIEQALEAGAVVLIENLEESIDPVLGPLLGRE VIKKGRFIKIGDKECEYNPKFRLILHTKLANPHYQ PELQAQATLINFTVTRDGLEDQLLAAVVSMERP DLEQLKSDLTKQQNGFKITLKTLEDSLLSRLSSAS GNFLGETVLVENLEITKQTAAEVEKKVQEAKVT EVKINEAREHYRPAAARASLLYFIMNDLSKIHPM YQFSLKAFSIVFQKAVERAAPDESLRERVANLID SITFSVYQYTIRGLFECDKLTYLAQLTFQILLMNR EVNAVELDFLLRSPVQTGTASPVEFLSHQAWGA VKVLSSMEEFSNLDRDIEGSAKSWKKFVESECPE KEKLPQEWKNKTALQRLCMLRAMRPDRMTYAL RDFVEEKLGSKYVVGRALDFATSFEESGPATPMF FILSPGVDPLKDVESQGRKLGYTFNNQNFHNVSL GQGQEVVAEAALDLAAKKGHWVILQNTLEMCS RETEFKSILFALCYFHAVVAERRKFGPQGWNRSY PFNTGDLTISVNVLYNFLEANAKVPYDDLRYLFG EIMYGGHITDDWDRRLCRTYLGEFIRPEMLEGEL SLAPGFPLPGNMDYNGYHQYIDAELPPESPYLYG LHPNAEIGFLTQTSEKLFRTVLELQPRDSQARDG AGATREEKVKALLEEILERVTDEFNIPELMAKVE ERTPYIVVAFQECGRMNILTREIQRSLRELELGLK GELTMTSHMENLQNALYFDMVPESWARRAYPS TAGLAAWFPDLLNRIKELEAWTGDFTMPSTVWL TGFFNPQSFLTAIMQSTARKNEWPLDQMALQCD MTKKNREEFRSPPREGAYIHGLFMEGACWDTQA GIITEAKLKDLTPPMPVMFIKAIPAD\RQDCGHVY SCPVTKTSQ\RDPTYVWTFNLKTKENPSKWVLA GVALLLQI
3509	A	3	6388	ILYINPADLGWNPPVSSWIEKREIQTERANLTILF DKYLPTCLDTLRTRFKKIIPIPEQSMVQMVCHLLE CLLTTEDIPADCPKEIYEHYFVFAAIWAFGGAMV QDQLVDYRAEFSKWWLTEFKTVKFPSQGTIFDY YIDPETKKFEPWSKLVPQFEFDPEMPLQACLVHT SETIRVCYFMERLMARQRPVMLVGTAGTGKSVL VGAKLASLDPEAYLVKNVPFNYYTTSAMLQAVL EKPLEKKAGRNYGPPGNKKLIYFIDDMNMPEVD AYGTVQPHTIIRQHLDYGHWYDRSKLSLKEITNV QYVSCMNPTAGSFTINPRLQRHFSVFVLSFPGAD ALSSIYSIILTQHLKLGNFPASLQKSIPPLIDLALAF HQKIATTFLPTGIKFHYIFNLRDFANIFQGILFSSV ECVKSTWDLIRLYLHESNRVYRDKMVEEKDFDL FDKIQTEVLKKTFDDIEDPVEQTQSPNLYCHFAN GIGEPKYMPVQSWELLTQTLVEALENHNEVNTV MDLVLFEDAMRHVCHINRILESPRGNALLVGVG GSGKQSLTRLAAFISSMDVFQITLRKGYQIQDFK MDLASLCLKAGVKNLNTVFLMTDAQVADERFL VLINDLLASGEIPDLYSDDEVENIISNVRNEVKSQ GLVDNRENCWKFFIDRIRRQLKVTLCFSPVGNKL RVRSRKFPAIVNCTAIHWFHEWPQQALESVSLRF LQNTEGIEPTVKQSISKFMAFVHTSVNQTSQSYLS NEQRYNYTTPKSFLEFIRLYQSLLHRHRKELKCK TERLENGLLKLHSTSAQVDDLKAKLAAQEVELK QKNEDADKLIQVVGVETDKVSREKAMADEEEQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KVAVIMLEVKQKQKDCEEDLAKAEPALTAAQA ALNTLNKTNLTELKSFGSPPLAVSNVSAAVMVL MAPRGRVPKDRSWKAAKVTMAKVDGFLDSLIN FNKENIHENCLKAIRPYLQDPEFNPEFVATKSYA AAGLCSWVINIVRFYEVFCDVEPKRQALNKATA DLTAAQEKLAAIKAKIAHLNENLAKLTARFEKA TADKLKCQQEAEVTAVTISLANRLVGGLASENV RWADAVQNFKQQERTLCGDILLITAFISYLGFFT KKYRQSLLDRTWRPYLSQLKTPIPVTPALDPLRM LMDDADVAAWQNEGLPADRMSVENATILINCE RWPLMVDPQLQGIKWIKNKYGEDLRVTQIGQKG YLQIIEQALEAGAVVLIENLEESIDPVLGPLLGRE VIKKGRFIKIGDKECEYNPKFRLILHTKLANPHYQ PELQAQATLINFTVTRDGLEDQLLAAVVSMERP DLEQLKSDLTKQQNGFKITLKTLEDSLLSRLSSAS GNFLGETVLVENLEITKQTAAEVEKKVQEAKVT EVKINEAREHYRPAAARASLLYFIMNDLSKIHPM YQFSLKAFSIVFQKAVERAAPDESLRERVANLID SITFSVYQYTIRGLFECDKLTYLAQLTFQILLMNR EVNAVELDFLLRSPVQTGTASPVEFLSHQAWGA VKVLSSMEEFSNLDRDIEGSAKSWKKFVESECPE KEKLPQEWKNKTALQRLCMLRAMRPDRMTYAL RDFVEEKLGSKYVVGRALDFATSFEESGPATPMF FILSPGVDPLKDVESQGRKLGYTFNNQNFHNVSL GQGQEVVAEAALDLAAKKGHWVILQNTLEMCS RETEFKSILFALCYFHAVVAERRKFGPQGWNRSY PFNTGDLTISVNVLYNFLEANAKVPYDDLRYLFG EIMYGGHITDDWDRRLCRTYLGEFIRPEMLEGEL SLAPGFPLPGNMDYNGYHQYIDAELPPESPYLYG LHPNAEIGFLTQTSEKLFRTVLELQPRDSQARDG AGATREEKVKALLEEILERVTDEFNIPELMAKVE ERTPYIVVAFQECGRMNILTREIQRSLRELELGLK GELTMTSHMENLQNALYFDMVPESWARRAYPS TAGLAAWFPDLLNRIKELEAWTGDFTMPSTVWL TGFFNPQSFLTAIMQSTARKNEWPLDQMALQCD MTKKNREEFRSPPREGAYIHGLFMEGACWDTQA GIITEAKLKDLTPPMPVMFIKAIPAD\RQDCGHVY SCPVTKTSQ\RDPTYVWTFNLKTKENPSKWVLA GVALLLQI
3510	A	390	3330	AAGSGSRPPAPAARKMADLAECNIKVMCRFRPL NESEVNRGDKYIAKFQGEDTVVIASKPYAFDRVF QSSTSQEQVYNDCAKKIVKDVLEGYNGTIFAYG QTSSGKTHTMEGKLHDPEGMGIIPRIVQDIFNYIY SMDENLEFHIKVSYFEIYLDKIRDLLDVSKTNLSV HEDKNRVPYVKGCTERFVCSPDEVMDTIDEGKS NRHVAVTNMNEHSSRSHSIFLINVKQENTQTEQK LSGKLYLVDLAGSEKVSKTGAEGAVLDEAKNIN KSLSALGNVISALAEGSTYVPYRDSKMTRILQDS LGGNCRTTIVICCSPSSYNESETKSTLLFGQRAKTI KNTVCVNVELTAEQWKKKYEKEKEKNKILRNTI QWLENELNRWRNGETVPIDEQFDKEKANLEAFT VDKDITLTNDKPATAIGVIGNFTDAERRKCEEEIA KLYKQLDDKDEEINQQSQLVEKLKTQMLDQEEL LASTRRDQDNMQAELNRLQAENDASKEEVKEV LQALEELAVNYDQKSQEVEDKTKEYELLSDELN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QKSATLASIDAELQKLKEMTNHQKKRAAEMMA SLLKDLAEIGIAVGNNDVKQPEGTGMIDEEFTVA RLYISKMKSEVKTMVKRCKQLESTQTESNKKME ENEKELAACQLRISQHEAKIKSLTEYLQNVEQKK RQLEESVDALSEELVQLRAQEKVHEMEKEHLNK VQTANEVKQAVEQQIQSHRETHQKQISSLRDEVE AKAKLITDLQDQNQKMMLEQERLRVEHEKLKA TDQEKSRKLHELTVMQDRREQARQDLKGLEETV AKELQTLHNLRKLFVQDLATRVKKSAEIDS\DDT GGSAAQKQKISFLENNLE\QLTKSAQTSWYRDNA DLRCELPKLEKRLRATAERVKALESALKEAKEN ASRDRKRYQQEVDRIKEAVRSKNMARRGHSAQI AKPIRPGQHPAASPTHPSAIRGGGAFVQNSQPVA VRGGGGKQV
3511	A	1	1757	MASVQASRRQWCYLCDLPKMPWAMVWDFSEA VCRGCVNFEGADRIELLIDAARQLKRSHVLPEGR SPGPPALKHPATKDLAAAAAQGPQLPPPQAQPQP SGTGGGVSGQDRYDRATSSGRLPLPSPALEYTLG SRLANGLGREEAVAEGARRALLGSMPGLMPPGL LAAAVSGLGSRGLTLAPGLSPARPLFGSDFEKEK QQRNADCLAELNEAMRGRAEEWHGRPKAVREQ LLALSACAPFNVRFKKDHGLVGRVFAFDATARP PGYEFELKLFTEYPCGSGNVYAGVLAVARQMFH DALREPGKALASSGFKYLEYERRHGSGEWRQLG ELLTDGVRSFREPAPAEALPQQYPEPAPAALCGP PPRAPSRNLAPTPRRRKASPEPEGEAAGKMTTEE QQQRHWVAPGGPYSAETPGVPSPIAALKNVAEA LGHSPKDPGGGGGPVRAGGASPAASSTAQPPTQ HRLVARNGEAEVSPTAGAEAVSGGGSGTGATPG APLC\CTLCRERLEDTHFVQ\CPPVPEHKFCFPCSR KFIKAQGPAGE\VYCPSGDKCPLVGSSVPWAFMQ GEIATILAGDIKVKKERDP
3512	A	3	1994	NTNSSSVTNSAAGVEDLNIVQVTVPDNEKERLSS IEKIKQLREQVNDLFSRKFGEAIGVDFPVKVPYR KITFNPGCVVIDGMPPGVVFKAPGYLEISSMRRIL EAAEFIKFTVIRPLPGLELSNGEYSTVGKRKIDQE GRVFQEKWERAYFFVEVQNISTCLICKRSMSVSK EYNLRRHYQTNHSKHYDQYMERMRDEKLHELK KGLRKYLLGLSDTECPEQKQVFANPSPTQKSPVQ PVEDLAGNLWEKLREKIRSFVAYSIAIDEITDINN TTQLAIFIRGVDENFDVSEELLDTVPMTGTKSGN EIFSRVEKSLKNFCINWSKLVSVASTGTPPMVDA NNGLVTKLKSRVATFCKGAELKSICCIIHPESLCA Q\KLKMDHVMDVVVKSVNWICSRGLNHSEFTTL LYELDSQYGSLLYYTEIKWLSRGLVLKRFFESLE EIDSFMSSRGKPLPQLSSIDWIRDLAFLVDMTMH LNALNISLQGHSQIVTQMYDLIRAFLAKLCLWET HLTRNNLAHFPTLKLVSRNESDGLNYIPKIAELK TEFQKRLSDFKLYESELTLFSSPFSTKIDSVHEELQ MEVIDLQCNTVLKTKYDKVGIPEFYKYLWGSYP KYKHHCAKILSMFGSTYICEQLFSIMKLSKTKYC SQLKDSQWDSVLHIAT
3513	A	1836	513	FKSLLSVKWFCFSILVLIFLGTRCYWEMTQSRPSP DPHRGRWEGGRSRPKGGEEGRRRTRVPGLVTAS GPGNPLPDRLGEMAGGRHRRVVGTLHLLLLVAA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LPWASRGVSPSASAWPEEKNYHQPAILNSSALRQ IAEGTSISEMWQNDLQPLLIERYPGSPGSYAARQ HIMQRIQRLQADWVLEIDTFLSQTPYGYRSFSNII STLNPTAKRHLVLACHYDSKYFSHW\NNRVFVG ATDSAVPCAMMLELARALDKKLLSLKTVSDSKP DLSLQLIFFDGEEAFLHWSPQDSLYGSRHLAAKM ASTPHPPGARGTSQLHGMDLLVLLDLIGAPNPTF PNFFPNSARWFERLQAIEHELHELGLLKDHSLEG RYFQNYSYGGVIQDDHIPFLRRGVPVLHLIPSPFP EVWHTMDDNEENLDESTIDNLNKILQVFVLEYL HL
3514	A	1836	513	FKSLLSVKWFCFSILVLIFLGTRCYWEMTQSRPSP DPHRGRWEGGRSRPKGGEEGRRRTRVPGLVTAS GPGNPLPDRLGEMAGGRHRRVVGTLHLLLLVAA LPWASRGVSPSASAWPEEKNYHQPAILNSSALRQ IAEGTSISEMWQNDLQPLLIERYPGSPGSYAARQ HIMQRIQRLQADWVLEIDTFLSQTPYGYRSFSNII STLNPTAKRHLVLACHYDSKYFSHW\NNRVFVG ATDSAVPCAMMLELARALDKKLLSLKTVSDSKP DLSLQLIFFDGEEAFLHWSPQDSLYGSRHLAAKM ASTPHPPGARGTSQLHGMDLLVLLDLIGAPNPTF PNFFPNSARWFERLQAIEHELHELGLLKDHSLEG RYFQNYSYGGVIQDDHIPFLRRGVPVLHLIPSPFP EVWHTMDDNEENLDESTIDNLNKILQVFVLEYL HL
3515	A	114	754	LCRDLTTTMSSKRTKTKTKKRPQRATSNVFAMF DQSQIQEFKEAFNMIDQNRDGFIDKEDLHDMLAS LGKNPTDEYLDAMMNEAPGPINFTMFLTMFGEK LNGTDPEDVIRNAFACFDEEATGTIQEDYLRELL TT\MGDRF\TDE\EVDELYREAPI\DKKGGIFNYI\E FTRHLETGGPKDKDDRKITFQIPSPNVPWLATFG VFLEIFLLHGP
3516	A	1	5169	MAAAPSALLLLPPFPVLSTYRLQSRSRPSAPETDD SRVGGIMRGEKNYYFRGAAGDHGSCPTTTSPLA SALLMPSEAVSSSWSESGGGLSGGDEEDTRLLQL LRTARDPSEAFQALQAALPRRGGRLGFPRRKEAL YRALGRVLVEGGSDEKRLCLQLLSDVLRGQGEA GQLEEAFSLALLPQLVVSLREENPALRKDALQIL HICLKRSPGEVLRTLIQQGLESTDARLRASTALLL PILLTTEDLLLGLDLTEVIISLARKLGDQETEEESE TAFSALQQIGERLGQDRFQSYISRLPSALRRHYN RRLESQFGSQVPYYLELEASGFPEDPLPCAVTLS NSNLKFGIIPQELHSRLLDQEDYKNRTQAVEELK QVLGKFNPSSTPHSSLVGFISLLYNLLDDSNFKVV HGTLEVLHLLVIRLGEQVQQFLGPVIAASVKVLA DNKLVIKQEYMKIFLKLMKEVGPQQVLCLLLEH LKHKHSRVREEVVNICICSLLTYPSEDFDLPKLSF DLAPALVDSKRRVRQAALEAFAVLASSMGSGKT SILFKAVDTVELQDNGDGVMNAVQARLARKTLP RLTEQGFVEYAVLMPSSAGGRSNHLAHGADTD WLLAGNRTQSAHCHCGDHVRDSMHIYGSYSPTI CTRRVLSAGKGKNKLPWENEQPGIMGENQTSTS KDIEQFSTYDFIPSAKLKLSQGMPVNDDLCFSRK RVSRNLFQNSRDFNPDCLPLCAAGTTGTHQTNLS GKCAQLGFSQICGKTGSVGSDLQFLGTTSSHQEK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VYASLNFGSKTQQTFGSQTECTSSNGQNPSPGAYILPSYPVSSPRTSPKHTSPLIISPKKSQDNSVNFSNSWPLKSFEGLSKPKSHRRSLSAQKSS\DPTGR\NHG\ENSQEKPP\VQLTPAL\VRSPSSRRGLNGTKPVPPIP\RGISLLPDKADLSTVGHKKKEPDDIWKCEKDSLPIDLSELNFKDKDLDQEEMHSSLRSLRNSAAKKRAKLSGSTSDLESPDSAMKLDLTMDSPSLSSSPNINSYSESGVYSQESLTSSLSTTPQGKRIMSDIFPTFGSKPCPTRLSSAKKKISHIAEQSPSAGSSSNPQQISSFDFTTTKALSEDSVVVVGKGVFGSLSSAPATCSQSVISSVENGDTFSIKQSIEPPSGIYGRSVQQNISSYLDVENEKDAKVSISKSTYNKMRQKRKEEKELFHNKDCEKKEKNSWERMRHTGTEKMASESETPTGAISQYKERMPSVTHSPEIMDLSELRPFSKPEIALTEALRLLADEDWEKKIEGLNFIRCLAAFHSEILNTKLHETNFAVVQEVKNLRSGVSRAAVVCLSDLFTYLKKSMDQELDTTVKVLLHKAGESNTFIREDVDKALRAMVNNVTPARAVVSLINGGQRYYGRKMLFFMMCHPNFEKMLEKYVPSKDLPYIKDSVRNLQQKGLGEIPLDTPSAKGRRSHTGSVGNTRSSSVSRDAFNSAERAVTEVREVTRKSVPRNSLESAEYLKLITGLLNAKDFRDRINGIKQLLSDTENNQDLVVGNIVKIFDAFKSRLHDSNSKVNLVALETMHKMIPLLRDHLSPIINMLIPAIVDNNLNSKNPGIYAAATNVVQALSQHVDNYLLLQPFCTKAQFLNGKAKQDMTEKLADIVTELYQRKPHATEQKVLVVLWHLLGNMTNSGSLPGAGGNIRTATAKLSKALFAQMGQNLLNQAASQPPHIKKSLEELLDMTILNEL
3517	A	1449	252	QDLKPVLDREYLAIYLKMVFFTCNACGESVKKIQVEKHVSVCRNCECLSCIDCGKDFWGDDYKNHVKCISEDQKYGGKGY/EKVKTHKGD/ASKQQAWIQKISELIK\RPNVSPKVRELLEQISAFDNVPQ\KKAKFQNWMKNSLKVHNESILDQVWNIFSEASNSEPVNKEQDQRPLHPVANPHAEISTKVPASKVKDAVEQQGEVKKNKRERKEERQKKRKREKKELKLENHQENSRNQKPKKRKKGQEADLEAGGEEVPEANGSAGKRSKKKKQRKDSASEEEARVGAGKRKRRHSKVETDSKKKKMKLPEHPEGGEPEDDEAPAKGKFNWKGTIKAILKQAPDNEITIKKLRKKVLAQYYTVTDEHHRSEEELLVIFNKKISKNPTFKLLKDKVKLVK
3518	A	3	635	APDSNARNDHFDACSLRVQAGLSSAGPALGNSGLAALMASPSKAVIVPGNGGGDVTTHGWYGWVKKELEKIPGFQCLAKNMPDPITARESIWLPFMETELHCDEKTIIIGHSSGAIAAMRYAETHRVYAIVLVSAYTSDLGDENERASGYFTRPWQWEKIKANCPYIVQFGSTDDPFLPWKEQQEVAD\SWKPNCTNSLTVATFRTQSFMN
3519	A	81	2277	VRETRREMAMAMSDSGASRLRRQLESGGFEARLYVKQLSQQSDGDRDLQEHRQRIQALAEETAQNLKRNVYQNYRQFIETAREISYLESEMYQLSHLLTEQKSSLESIPLTLLPAAAAAGAAAASGGEEGVGGAGGRDHLRGQAGFFSTPGGASRDGSGPGEEGKQRTLTTLLEKVEGCRHLLETPGQYLVYNGDLVEYDADHMAQLQRVHGFLMNDCLLVATWLPQRRGM

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YRYNALYSLDGLAVVNVKDNPPMKDMFKLLMF PENRIFQAENAKIKREWLEVLEDTKRALSEKRRR EQEEAAAPRGPPQVTSKATNPFEDDEEEEPAVPE VEEEKVDLSMEWIQELPEDLDVCIAQRDFEGAV DLLDKLNHYLEDKPSPPPVKELRAKVEERVRQL TEVLVFELSPDRSLRGGPKATRRAVSQLIRLGQC TKACELFLRNRAAAVHTAIRQLRIEGATLLYIHK LCHVFFTSLLETAREFEIDFAGTDSGCYSAFVVW ARSAMGMFVDAFSKQVFDSKESLSTAAECVKVA KEHCQQLGDIGLDLTFIIHALLVKDIQGALHSYK EIIIEATKHRNSEEMWRRMNLMTPEALGKLKEE MKSCGVSNFEQYTGDDCWVNLSYTVVAFTKQT MGFLEEALKLYFPELHMVLLESLVEIILVAVQHV DYSLRCEQDPEKKAFIRQNASFLYETVL\PVVEK RFEEGVGKPAKQLQDLRNASRLIRVNPESTTSVV
3520	A	1706	540	FVAHLAWPWRADGDMEDGVLNEGFLVKRGHIV HNWKARWFILRQNTLVYYKLEGGRRVTPPKGRI LLDGCTITCPCLEYENRPLLIKLKTQTSTEYFLEA CSREE/RRDAWAFE\ITGAIHAGQARGKVQQLHS LRNSFKLPPHISLHRIVDKMHDSNTGIRSSPNMEQ GSTYKKTFLGSSLVDWLISNSFTASRLEAVTLAS MLMEENFLRPVGVRSMGAIRSGDLAEQFLDDST ALYTFAESYKKKISPKEEISLSTVELSGTVVKQGY LAKQGHKRKNWKVRRFVLRKDPAFLHYYDPSK EENRPVGGFSLRGSLVSALEDNGVPTGVKGNVQ GNLFKVITK\DDTHYYIQA\SSKAE\RAE\WIGSLS KSLNMNKDPEGTPDSLPSLPR
3521	A	3	3063	HASVSLSLGCPRPCADTPGPQPQPMDLRVGQRPP VEPPPEPTLLALQRPQRLHHHLFLAGLQQQRSVE PMRVKMELPACGATLSLVPSLPAFSIPRHQSQSST PCPFLGCRPCPQLSMDTPMPELQEAPQEQELRQL LHKDKSKRSAVASSVVKQKLAEVILKKQQAALE RTVHPNSPGIPYRTLEPLETEGATRSMLSSFLPPV PSLPSDPPEHFPLRKTVSEPNLKLRYKPKKSLERR KNPLLRKESAPPSLRRRPAETLGDSSPSSSSTPAS GCSSPNDSEHGPNPILGSEALLGQRLRLQETSVAP FALPTVSLLPAITLGLPAPARADSDRRTHPTLGPR GPILGSPHTPLFLPHGLEPEAGGTLPSRLQPILLLD PSGSHAPLLTVPGLGPLPFHFAQSLMTTERLSGSG LHWPLSRTRSEPLPPSATAPPPPGPMQPRLEQLKT HVQVIKRSAKPSEKPRLRQIPSAEDLETDGGGPG QVVDDGLEHRELGHGQPEARGPAPLQQHPQVLL WEQQRLAGRLPRGSTGDTVLLPLAQGGHRPLSR AQSSPAAPASLSAPEPASQARVLSSSETPARTLPF TTGLIYDSVMLKHQCSCGDNSRHPEHAGRIQSIW SRLQERGLRSQCECLRGRKASLEELQSVHSERHV LLYGTNPLSRLKLDNGKLAGLLAQRMFVMLPCG GVGVDTDTIWNELHSSNAARWAAGSVTDLAFK VASRELKNGFAVVRPPGHHADHSTAMGFCFFNS VAIACRQLQQQSKASKILIVDWDVHHGNGTQQT FYQDPSVLYISLHRHDDGNFFPGSGAVDEVGAGS GEGFNVNVAWAGGLDPPMGDPEYLAAFRIVVM PIAREFSPDLVLVSAGFDAAEGHPAPLGGYHVSA KCFGYMTQQLMNLAGGAVVLALEGGHDLTAIC DASEACVAALLGNRVDPLSEEGWKQKPNLNAIR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SLEA\VIRVHSKYWGCMQRLASCPDSWVPRVPG ADKEEVEAVTALASLSVGILAEDRPSEQLVEEEE PMNL
3522	A	9	602	KMAALGEPVRLERDICRAIELLEKLQRSGEVPPQ KLQALQRVLQSEFCNAVREVYEHVYETVDISSSP EVRANATAKATVAAFAASEGHSHPRVVELPKTE EGLGFNIMGGKEQNSPIYISRIIP/GGIADRHGGLK RGDQLLSVNGVSVEGEHHEKAVELLKAAQGKV KLVVRYTPKVLEEMESRFEKMRSAKRRQQT
3523	A	645	1465	IMAETSLLEAGASAASTAAALENLQVEASCSVCL EYLKEPVIIECGHNFCKACITRWWEDLERDFPCP VCRKTSRYRSLRPNRQLGSMVEIAKQL\RPSSGRS GMRASAPQHHEALSLFCYEDQEAVCLICAISHTH RAHTVVPLDDATQEYKEKLQKCLEA\LNQKLQEI TRCKSSEEKKPGELKRLVESRRQQILREFEELHRR LDEEQQVLLSRLEEEEQDILQRLRENAAHLGDKR RDLAHLAAEVEGKCLQSGFEMLKVRPLPLHSPS G
3524	A	3	698	PMVRHEAGEALGAIGDPEVLEILKQYSSDPVIEV AETCQLAVRRLEWLQQHGGEPAAGPYLSVDPAP PAEER\DVGRLREALLDESRPLFERYRAMFALRN AGGEEAALALAEGLHCGSALFRHEVGYVLGQLQ HEAAVPQLAAALARCTENPMVRHECAEALGAIA RPACLAALQAHADDPERVVRE\SCKVALDMYEH ETGRAFQYADGLEQLRGAPSLGPNPHPELPEDS
3525	A	1452	694	EGLQRPEYLVASAAGFQGLAWGGEGRGRAGCS SSGFRDAEPLLLSCPGRNEPLKKERLKWKSDYP MTDGQLRSKRDEFWDTAPAFEGRKEIWDALKA AAYAAEANDHELAQAILDGASITLPHGTLCECY DELGNRYQLPIYCLSPPVNLLLEHTEEESLEPPEP PPSVRREFPLKVRLSTGKDVRLSASLPDTVGQLK RQLHAQE/GTPKPSWQRWFFSGKLLTDRTRLQET KIQKDFVIQVIINQPPPPQD
3526	A	123	3441	PGNEGLGLAADHNEDLGHLSADAPWPAVTMAP RKRSHHGLGFLCCFGGSDIPEINLRDNHPLQFME FSSPIPNAEELNIRFAELVDELDLTDKNREAMFAL PPEKKWQIYCSKKKEQEDPNKLATSWPDYYIDRI NSMAAMQSLYAFDEEETEMRNQVVEDLKTALR TQPMRFVTRFIELEGLTCLLNFLRSMDHATCESRI HTSLIGCIIALMNNSQGRAHVLAQPEAISTIAQSL RTENSKTKVAVLEILGAVCLVPGGHKKVLQAML HYQVYAAERTRFQTLLNELDRSLGRYRDEVNLK TAIMSFINAVLNAGAGEDNLEFRLHLRYEFLMLG IQPVIDKLRQHENAILDKHLDFFEMVRNEDDLEL ARRFDMVHIDTKSASQMFELIHKKLKYTEAYPC LLSVLHHCLQMPYKRNGGYFQQWQLLDRILQQI VLQDERGVDPDLAPLENFNVKNIVNMLINENEV KQWRDQAEKFRKEHMELVSRLERKERECETKTL EKEEMMRT\LNKMKDKLARESQELRQARGQVA ELVAQLSELSTGPVSSPPPPGGPLTLSSSMTTNDL PPPPPPLPFACCPPPPPPPLPPGGPPTPPGAPPCLG MGLPLPQDPYPSSDVPLRKKRVPQPSHPLKSFNW VKLNEERVPGTVWNEIDDMQVFRILDLEDFEKM FSAYQRHQELITNPSQQKELGSTEDIYLASRKVK ELSVIDGRRAQNCIILLSKLKLSNEEIRQAILKMD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EQEDLAKDMLEQLLKFIPEKSDIDLLEEHKHEIER MARADRFLYEMSRIDHYQQRLQALFFKKKFQER LAEAKPKVEAILLASRELVRSKRLRQMLEVILAI GNFMNKGQRGGAYGFRVASLNKIADTKSSIDRN ISLLHYLIMILEKHFPDILNMPSELQHLPEAAKVN LAELEKEVGNLRRGLRAVEVELEYQRRQVREPS DKFVPVMSDFITVSSFSFSELEDQLNEARDKFAK ALMHFGEHDSKMQPDEFFGIFDTFLQAFSEARQD LEAMRRRKEEEERRARMEAMLKEQRERERWQR QRKVLAAGSSLEEGGEFDDLVSALRSGEVFDKD LCKLKRSRKRSGSQALEVTRERAINRLNY
3527	A	1445	714	LLGTRMLAGQLEARDPKEGTHPEDPCPGAGAV MEKTAVAAEVLTEDCNTGEMPPLQQQIIRLHQE LGRQKSLWADVHGKLRSHIDALREQNMELREKL RALQLQRWKARKKSAASPHAGQESHTLALEPAF GKISPLSADEETIPKYAGHKN\QSGHSSWGQRSSS NNSAPPKPMSLKIERISSWKTPPQENRDKNLSRR RQDRRATPTGRPTPCAERRG\VSEDGKVASDTCV TLHWPLGKFRFR
3528	A	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLLTVK GLLKPSFSPRNYKALSEVQGWKQRMAAKELAR QNMDLGFKLLKKLAFYNPGRNIFLSPLSISTAFS MLCLGAQDSTLDEIKQGFNFRKMPEKDLHEGFH YIIHELTQKTQDLKLSIGNTLFIDQRLQPQRKFLE DAKNFYSAETILTNFQNLEMAQKQINDFI/ESKTH GKINNLIENIDPGTVMLLANYIFFRARWKHEFDP NVTKEEDFFLEKNSSVKVPMMFRSGIYQVGYDD KLSCTILEIPYQKNITAIFILPDEGKLKHLEKGLQV DTFSRWKTLLSRRVVDVSVPRLHMTGTFDLKKT LSYIGVSKIFEEHGDLTKIAPHRSLKVGEAVNKA ELKMDERGTEGAAGTGAQTLPMETPLVVKIDKP YLLLIYSEKIPSVLFLGKIVNPIGK
3529	A	1	5684	VSSVSHENPTEVFEDGENPPSSRSSESGFTEFIQY QADRTDDIDRELSEGQGAAAIPIGSTSSETETAST VGSEETIIQTPSVVTQGTATRSRKTAQKTAMQCC LEYVQQFLTRLINLYIIQNNSFSQSLATEHQGDLG REQGETSKWDRNSQGDVKEKNISKQKTSKEYLS AFLAACQLFLECSSFPVYIAEGNHTSELRSEKLET DCEHVQPPQWLQTLMNACSQASDFSVQSVAISL VMDLVGLTQSVAMVTGENINSVEPAQPLSPNQG RVAVVIRPPLTQGNLRYIAEKTEFFKHVALTLWD QLGDGTPQHHQKSVELFYQLHNLVPSSSICEDVI SQQLTHKDKKIRMEAHAKFAVLWHLTRDLHINK SSSFVRSFDRSLFIMLDSLNSLDGSTSSVGQAWL NQVLQRHDIARVLEPLLLLLLHPKTQRVSVQRV QAERYWNKSPCYPGEESDKHFMQNFACSNVSQ VQLITSKGNGEKPLTMDEIENFSLTVNPLSDRLSL LSTSSETIPMVVSDFDLPDQQIEILQSSDSGCSQSS AGDNLSYEVDPETVNAQEDSQMPKESSPDDDVQ QVVFDLICKVVSGLEVESASVTSQLEIEAMPPKC SDIDPDEETIKIEDDSIQQSQNALLSNESSQFLSVS AEGGHECVANGISRNSSSPCISGTTHTLHDSSVAS IETKSRQRSHSSIQFSFKEKLSEKVSEKETIVKESG KQPGAKPKVKLARKKDDDKKSSNEKLKQTSV FFSDGLDLENWYSCGEGDISEIESDMGSPGSRKSP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NFNIHPLYQHVLLYLQLYDSSRTLYAFSAIKAILK TNPIAFVNAISTTSVNNAYTPQLSLLQNLLARHRI SVMGKDFYSHIPVDSNHNFRSSMYIEILISLCLYY MRSHYPTHVKVTAQDLIGNRNMQMMSIEILTLL FTELAKVIESSAKGFPSFISDMLSKCKVQKVILHC LLSSIFSAQKWHSEKMAGKNLVAVEEGFSEDSLI NFSEDEFDNGSTLQSQLLKVLQRLIV\LEHRVM\T IPEE\NETGFDFVVS\DLEHISPHQPMTSLQYLHAQ SITCQGMFLCAVIRA\LHQHCACKMHPQWIGLIT STLPYMGKVLQRVVVSVTLQLCRNLDNLIQQYK YETGLSDSRPLWMASIIPPDMILTLLEGITAIIHYC LLDPTTQYHQLLVSVDQKHLFEARSGILSILHMI MSSVTLLWSILHQADSSEKMTIAASASLTTINLG ATKNLRQQILELLGPISMNHGVHFMAAIAFVWN ERRQNKTTTRTKVIPAASEEQLLLVELVRSISVM RAETVIQTVKEVLKQPPAIAKDKKHLSLEVCML QFFYAYIQRIPVPNLVDSWASLLILLKDSIQLSLP APGQFLILGVLNEFIMKNPSLENKKDQRDLQDVT HKIVDAIGAIAGSSLEQTTWLRRNLEVKPSPKIM VDGTNLESDVEDMLSPAMETANITPSVYSVHAL TLLSEVLAHLLDMVFYSDEKERVIPLLVNIMHYV VPYLRNHSAHNAPSYRACVQLLSSLSGYQYTRR AWKKEAFDLFMDPSFFQMDASCVNHWRAIMDN LMTHDKTTFRDLMTRVAVAQSSSLNLFANRDVE LEQRAMLLKRLAFAIFSSEIDQYQKYLPDIQERLV ESLRLPQVPTLHSQVFLFFRVLLLRMSPQHLTSL WPTMITELVQVFLLMEQELTADEDISRTSGPSVA GLETTYTGGNGFSTSYNSQRWLNLYLSACKFLD LALALPSENLPQFQMYRWAFIPEASDDSGLEVRR QGIHQREFKPYVVRLAKLLRKRAKKNPEEDNSG RTLGWEPGHLLLTICTVRSMEQLLPFFNVLSQVF NSKVTSRCGGHSGSPILYSNAFPNKDMKLENHKP CSSKARQKIEEMVEKDFLEGMIKT
3530	A	1	5684	VSSVSHENPTEVFEDGENPPSSRSSESGFTEFIQY QADRTDDIDRELSEGQGAAAIPIGSTSSETETAST VGSEETIIQTPSVVTQGTATRSRKTAQKTAMQCC LEYVQQFLTRLINLYIIQNNSFSQSLATEHQGDLG REQGETSKWDRNSQGDVKEKNISKQKTSKEYLS AFLAACQLFLECSSFPVYIAEGNHTSELRSEKLET DCEHVQPPQWLQTLMNACSQASDFSVQSVAISL VMDLVGLTQSVAMVTGENINSVEPAQPLSPNQG RVAVVIRPPLTQGNLRYIAEKTEFFKHVALTLWD QLGDGTPQHHQKSVELFYQLHNLVPSSSICEDVI SQQLTHKDKKIRMEAHAKFAVLWHLTRDLHINK SSSFVRSFDRSLFIMLDSLNSLDGSTSSVGQAWL NQVLQRHDIARVLEPLLLLLLHPKTQRVSVQRV QAERYWNKSPCYPGEESDKHFMQNFACSNVSQ VQLITSKGNGEKPLTMDEIENFSLTVNPLSDRLSL LSTSSETIPMVVSDFDLPDQQIEILQSSDSGCSQSS AGDNLSYEVDPETVNAQEDSQMPKESSPDDDVQ QVVFDLICKVVSGLEVESASVTSQLEIEAMPPKC SDIDPDEETIKIEDDSIQQSQNALLSNESSQFLSVS AEGGHECVANGISRNSSSPCISGTTHTLHDSSVAS IETKSRQRSHSSIQFSFKEKLSEKVSEKETIVKESG KQPGAKPKVKLARKKDDDKKKSSNEKLKQTSV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				FFSDGLDLENWYSCGEGDISEIESDMGSPGSRKSP NFNIHPLYQHVLLYLQLYDSSRTLYAFSAIKAILK TNPIAFVNAISTTSVNNAYTPQLSLLQNLLARHRI SVMGKDFYSHIPVDSNHNFRSSMYIEILISLCLYY MRSHYPTHVKVTAQDLIGNRNMQMMSIEILTLL FTELAKVIESSAKGFPSFISDMLSKCKVQKVILHC LLSSIFSAQKWHSEKMAGKNLVAVEEGFSEDSLI NFSEDEFDNGSTLQSQLLKVLQRLIV\LEHRVM\T IPEE\NETGFDFVVS\DLEHISPHQPMTSLQYLHAQ SITCQGMFLCAVIRA\LHQHCACKMHPQWIGLIT STLPYMGKVLQRVVVSVTLQLCRNLDNLIQQYK YETGLSDSRPLWMASIIPPDMILTLLEGITAIIHYC LLDPTTQYHQLLVSVDQKHLFEARSGILSILHMI MSSVTLLWSILHQADSSEKMTIAASASLTTINLG ATKNLRQQILELLGPISMNHGVHFMAAIAFVWN ERRQNKTTTRTKVIPAASEEQLLLVELVRSISVM RAETVIQTVKEVLKQPPAIAKDKKHLSLEVCML QFFYAYIQRIPVPNLVDSWASLLILLKDSIQLSLP APGQFLILGVLNEFIMKNPSLENKKDQRDLQDVT HKIVDAIGAIAGSSLEQTTWLRRNLEVKPSPKIM VDGTNLESDVEDMLSPAMETANITPSVYSVHAL TLLSEVLAHLLDMVFYSDEKERVIPLLVNIMHYV VPYLRNHSAHNAPSYRACVQLLSSLSGYQYTRR AWKKEAFDLFMDPSFFQMDASCVNHWRAIMDN LMTHDKTTFRDLMTRVAVAQSSSLNLFANRDVE LEQRAMLLKRLAFAIFSSEIDQYQKYLPDIQERLV ESLRLPQVPTLHSQVFLFFRVLLLRMSPQHLTSL WPTMITELVQVFLLMEQELTADEDISRTSGPSVA GLETTYTGGNGFSTSYNSQRWLNLYLSACKFLD LALALPSENLPQFQMYRWAFIPEASDDSGLEVRR QGIHQREFKPYVVRLAKLLRKRAKKNPEEDNSG RTLGWEPGHLLLTICTVRSMEQLLPFFNVLSQVF NSKVTSRCGGHSGSPILYSNAFPNKDMKLENHKP CSSKARQKIEEMVEKDFLEGMIKT
3531	A	553	2470	LISPSPALSSQDPALSLKENLEDISGWGLPEARSK ESVSFKDVAVDFTQEEWGQLDSPQRALYRDVM LENYQNLLALGPPLHKPDVISHLERGEEPWSMQ REVPRGPCPEWELKAVPSQQQGICKEEPAQEPIM ERPLGGAQAWGRQAGALQRSQAAP\GR\RTCHG LGRP\VEEFPLRCPLFAQQRVPEGGPLLDTRKNV QATEGRTKAPARLCAGENASTPSEPEKFPQVRRQ RGAGAGEGEFVCGECGKAFRQSSSLTLHRRWHS REKAYKCDECGKAFTWSTNLLEHRRIHTGEKPFF CGECGKAFSCHSSLNVHQRIHTGERPYKCSACEK AFSCSSLLSMHLRVHTGEKPYRCGECGKAFNQR THLTRHHRIHTGEKPYQCGSCGKAFTCHSSLTVH EKIHSGDKPFKCSDCEKAFNSRSRLTLHQRTHTG EKPFKCADCGKGFSCHAYLLVHRRIHSGEKPFKC NECGKAFSSHAYLIVHRRIHTGEKPFDCSQCWKA FSCHSSLIVHQRIHTGEKPYKCSECGRAFSQNHCL IKHQKIHSGEKSFKCEKCGEMFNWSSHLTEHQRL HSEGKPLAIQFNKHLLSTYYVPGSLLGAGDAGLR DVDPIDALDVAKLLCVVPPRAGRNFSLGSKPRN
3532	A	3931	317	HRELQDSPSAEPPAGSMPLRHWGMARGSKPVGD GAQPMAAMGGLKVLLHWAGPGGGEPWVTFSES

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SLTAEEVCIHIAHKVGITPPCFNLFALFDAQAQV WLPPNHILEIPRDASLMLYF\RHRFYSR\NWHGM NPREPAVYRCGPPGTEASSDQTAQGMQLLDPAS FEYLFEQGKHEFVNDVASLWELSTEEEIHHFKNE SLGMAFLHLCHLALRHGIPLEEVAKKTSFKDCIP RSFRRHIRQHSALTRLRLRNVFRRFLRDFQPGRLS QQMVMVKYLATLERLAPRFGTERVPVCHLRLLA QAEGEPCYIRDSGVAPTDPGPESAAGPPTHEVLV TGTGGIQWWPVEEEVNKEEGSSGSSGRNPQASL FGKKAKAHKAFGQPADRPREPLGAYFCDFRDIT HVGLKEHCVSIHRQDNKCLELSLPSRAAALSFVS LVDGYFRLTADSSHYLCHEVAPPRLVMSIRDGIH GPLLEPFVQAKLRPEDGLYLIHWSTSHPYRLILTV AQRSQAPDGMQSLRLRKFPIEQQDGAFVLEGWG RSFPSVRELGAALQGCLLRAGDDCFSLRRCCLPQ PGETSNLIIMRGARASPRTLNLSQLSFHRVDQKEI TQLSHLGQGTRTNVYEGRLRVEGSGDPEEGKMD DEDPLVPGRDRGQELRVVLKVLDPSHHDIALAF YETASLMSQVSHTHLAFVHGVCVRGPENIMVTE YVEHGPLDVWLRRERGHVPMAWKMVVAQQLA SALSYLENKNLVHGNVCGRNILLARLGLAEGTSP FIKLSDPGVGLGALSREERVERIPWLAPECLPGG ANSLSTAMDKWGFGATLLEICFDGEAPLQSRSPS EKEHFYQRQHRLPEPSCPQLATLTSQCLTYEPTQ RPSFRTILRDLTRLQPHNLADVLTVNPDSPASDPT VFHKRYLKKIRDLGEGHFGKVSLYCYDPTNDGT GEMVAVKALKADCGPQHRSGWKQEIDILRTLYH EHIIKYKGCCEDQGEKSLQLVMEYVPLGSLRDYL PRHSIGLAQLLLFAQQICEGMAYLHAQHYIHRDL AARNVLLDNDRLVKIGDFGLAKAVPEGHEYYRV REDGDSPVFWYAPECLKEYKFYYASDVWSFGVT LYELLTHCDSSQSPPTKFLELIGIAQGQMTVLRLT ELLERGERLPRPDKCPCEVYHLMKNCWETEASF RPTFENLIPILKTVHEKYQGQAPSVFSVC
3533	A	182	3465	FRWLDFFRGSINSQFEFGRKKENMTSPAKFKKDK EIIAEYDTQVKEIRAQLTEQMKCLDQQCELRVQL LQDLQDFFRKKAEIEMDYSRNLEKLAERFLAKT RSTKDQQFKKDQNVLSPVNCWNLLLNQVKRES RDHTTLSDIYLNNIIPRFVQVSEDSGRLFKKSKEV GQQLQDDLMKVLNELYSVMKTYHMYNADSISA QSKLKEAEKQEEKQIGKSVKQEDRQTPRSPDSTA NVRIEEKHVRRSSVKKIEKMKEKRQAKYTENKL KAIKARNEYLLALEATNASVFKYYIHDLSDLIDQ CCDLGYHASLNRALRTFLSAELNLEQSKHEGLD AIENAVENLDATSDKQRLMEMYNNVFCPPMKFE FQPHMGDMASQLCAQQPVQSELLQRCLQLQSRL STLKIENEEVKKTMEATLQTIQDIVTVEDFDVSD CFQYSNSMESVKSTVSETFMSKPSIAKRRANQQE TEQFYFTKMKEYLEGRNLITKLQAKHDLLQKTL GESQRTDCSLARRSSTVRKQDSSQAIPLVVESCIR FISRHGLQHEGIFRVSGSQVEVNDIKNAFERGEDP LAGDQNDHDMDSIAGVLKLYFRGLEHPLFPKDIF HDLMACVTMDNLQERALHIRKVLLVLPKTTLII MRYLFAFLNHLSQFSEENMMDPYNLAICFGPSL MSVPEGHDQVSCQAHVNELIKTIIIQHENIFPSPRE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LEGPVYSRGGSMEDYCDSPHGETTSVEDSTQDVTAEHHTSDDECEPIEAIAKFDYVGRTARELSFKKGASLLLYQRASDDWWEGRHNGIDGLIPHQYIVVQDTEDGVVERSSPKSEIEVISEPPEEKVTARAGASCPSGGHVADIYLANINKQRKRPESGSIRKTFRSDSHGLSSSLTDSSSPGVGASCRPSSQPIMSQSLPKEGPDKCSISGHGSLNSISRHSSLKNRLDSPQIRKTATAGRSKSFDNHRPMDPEVIAQDIEATMNSALNELRELERQSSVKHTPDVVLDTLEPLKTSPVVAPTSEPSSPLHTQLLKDPEPAFQRSASTAGDIACAFRPVKSVKMAAPVKPPAT\RPKPT\VFPKTNATSPGVNSSTSPQSTDKSCTV
3534	A	1	2640	FRRFVCPASRRPAAGLRDAASSAPRGMASEGPREPESEGIKLSADVKPFVPRFAGLNVAWLESSEACVFPSSAATYYPFVQEPPVTEQKIYTEDMAFGASTFPPQYLSSEITLHPYAYSPYTLDSTQNVYSVPGSQYLYNQPSCYRGFQTVKHRNENTCPLPQEMKALFKKKTYDEKKTYDQQKFDSERADGTISSEIKSARGSHHLSIYAENSLKSDGYHKRTDRKSRIIAKNVSTSKPEFEFTTLDFPELQGAENNMSEIQKQPKWGPVHSVSTDISLLREVVKPAAVLSKGEIVVKNNPNESVTANAATNSPSCTRELSWTPMGYVVRQTLSTELSAAPKNVTSMINLKTIASSADPKNVSIPSSEALSSDPSYNKEKHIIHPTQKSKASQGSDLEQNEASRKNKKKKEKSTSKYEVLTVQEPPRIEDAEEFPNLAVASERRDRIETPKFQSKQQPQDNFKNNVKKSQLPVQLDLGGMLTALEKKQHSQHAKQSSKPVVVSVGAVPVLSKECASGERGRRMSQMKTPHNPLDSSAPLMKKGKQREIPKAKKPTSLKKIILKERQERKQRLQENAVSPAFTSDDTQDGESGGDDQFPEQAELSGPEGMDELISTPSVEDKSEEPPGTELQRDTEASHLAPNHTTFPKIHSRRFRDYCSQMLSKEVDACVTDLLKELVRFQDRMYQKDPVKAKTKRRLVLGLREVLKHLKLKKLKCVIISPNCEKIQSKGGLDDTLHTIIDYACEQNIPFVFALNRKALGRSLNKAVPVSVVGIFSYDGAQDQFHKMVELTVAARQAYKTMLENVQQELVGEP\SLRHLPAYPHRAPAALQKMAPQP/VKEKEEPHYIEIWKKHLEAYSGCTLELEESLEASTSQMMNLNL
3535	A	1747	983	LFQFQVCRSVLSPRAAGCTWSLAPRSRGAAGSPRRYRGPQPQPAPPSALPNSRPSPVASGREMVVLSVPAEVTVILLDIEGTTTPIAFVKDILFPYIEENVKEYLQTHWEEEECQQDVSLLRKQV\FADVVPAVRKWREAGMKVYIYSSGSVEAQKLLFGHSTEGDILELVDGHFDTKIGHKVESESYRKIADSIGCSTNNILFLTDVTREASAAEEADVHVAVVVRPGNAGLTDDEKTYYSLITSFSELYLPSST
3536	A	3	1302	GRPPTAPHTGRPPTANRGDPRLDLKRGCARLLTSIESRGRPAASAGLRRDRCALRRWPLRRAPLARATRRRAGSPRRCAPRPRACPQGWSRARHQPGGLCLLLLLLCQFMEDRSAQAGNCWLRQAKNGRCQVLYKTELSKEECCSTGRLSTSWTEEDVNDNTLFKWMIFNGGAPNCIPCKETCENVDCGPGKKCRMNKKNKPRCVCAPDCSNITWKGPVCGLDGKTYRNECALLKARCKEQPELEVQYQGRCKKTCRDVFCPGSS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TCV\VDQTNNAYCVTCNRICPEPASSEQYLCGND GVTYS\SACHLRKATCLLGRSIGLAYEGKCIKAK SCEDIQCTGGKKCLWDFKVGRGRCSLCDELCPD SKSDEPVCASDNATYASECAMKEAACSSGVLLE VKHSGSCNSISEDTEEEEEDEDQDYSFPISSILEW
3537	A	285	2123	IGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRF LTSIPTGIPEDATTLYLQNNQINNAGIPSDLKNLL KVERIYLYHNSLDEFPTNLPKYVKELHLQENNIR TITYDSLSKIPYLEELHLDDNSVSAVSIEEGAFRD SNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRIS TISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFF NLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQ DNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQ GIFDDLDNITQLILRNNPWYCGCKMKWVRDWL QSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELF DCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTK QPDIKNPKLTKDHQTTGSPSRKTITITVKSVTSDTI HISWKLALPMTALRLSWLKLGHSPAFGSITETIVT GERSEYLVTALEPDSPYKVCMVPMETSNLYLFD ETPVCIETETAPLRMYNPTTTLNREQEKEPYKNP NLPLAAIIGGAVALVTIALLALVCWYVHRNGSLF SRNCAYSKGRRRKDDYAEAGTKKDNSILEIRETS FQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNH
3538	A	877	6184	WNVKPSLLVVQLFKFSDKEEHEQNDSISGKTGET GVEEMIATRKVEQDSKETVKLSHEDDHILEDAGS SDISSDAACTNPNKTENSLVGLPSCVDEVTECNL ELKDTMGIADKTENTLERNKIEPLGYCEDAESNR QLESTEFNKSNLEVVDTSTFGPESNILENAICDVP DQNSKQLNAIESTKIESHETANLQDDRNSQSSSV SYLESKSVKSKHTKPVIHSKQNMTTDAPKKIVAA KYEVIHSKTKVNVKSVKRNTDVPESQQNFHRPV KVRKKQIDKEPKIQSCNSGVKSVKNQAHSVLKK TLQDQTLVQIFKPLTHSLSDKSHAHPGCLKEPHH PAQTGHVSHSSQKQCHKPQQQAPAMKTNSHVK EELEHPGVEHFKEEDKLKLKKPEKNLQPRQRRSS KSFSLDEPPLFIPDNIATIRREGSDHSSSFESKYMW TPSKQCGFCKKPHGNRFMVGCGRCDDWFHGDC VGLSLSQAQQMGEEDKEYVCVKCCAEEDKKTEI LDPDTLENQATVEFHSGDKTMECEKLGLSKHTT NDRTKYIDDTVKHKVKILKRESGEGRNSSDCRD NEIKKWQLAPLRKMGQPVLPRRSSEEKSEKIPKE STTVTCTGEKASKPGTHEKQEMKKKKV\EKGVL NVHPAASASKPSADQIRQSVRHSLKDILMKRLTD SNLKVPEEKAAKVATKIEKELFSFFRDTDAKYKN KYRSLMFNLKDPKNNILFKKVLKGEVTPDHLIR MSPEELASKELAAWRRRENRHTIEMIEKEQREVE RRPITKITHKGEIEIESDAPMKEQEAAMEIQEPAA NKSLEKPEGSEK\RKEEVDSMSKDTTSQHRQHLF DLNCKICIGRMAPPVDDLSPKKVKVVVGVARKH SDNEAESIADALSSTSNILASEFFEEEKQESPKSTF SPAPRPEMPGTVEVESTFLARLNFIWKGFINMPS VAKFVTKAYPVSGSPEYLTEDLPDSIQVGGRISPQ TVWDYVEKIKASGTKEICVVRFTPVTEEDQISYT LLFAYFSSRKRYGVAANNMKQVKDMYLIPLGAT DKIPHPLVPFDGPGLELHRPNLLLGLIIRQKLKRQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HSACASTSHIAETPESAPPIALPPDKKSKIEVSTEEAPEEENDFFNSFTTVLHKQRNKPQQNLQEDLPTAVEPLMEVTKQEPPKPLRFLPGVLIGWENQPTTLELANKPLPVDDILQSLLGTTGQVYDQ\AQSVMEQNTVKEIPFLNEQTNSKIEKTDNVEVTDGENKEIKVKVDNISESTDKSAEIETSVVGSSSISAGSLTSLSLRGKPPDVSTEAFLTNLSIQSKQEETVESKEKTLKRQLQEDQENNLQDNQTSNSSPCRSNVGKGNIDGNVSCSENLVANTARSPQFINLKRDPRQAAGRSQPVTTSESKDGDSCRNGEKHMLPGLSHNKEHLTEQINVEEKLCSAEKNSCVQQSDNLKVAQNSPSVENIQTSQAEQAKPLQEDILMQNIETVHPFRRGSAVATSHFEVGNTCPSEFPSKSITFTSRSTSPRTSTNFSPMRPQQPNLQHLKSSPPGFPFPGPPNFPPQSMFGFPPHLPPPLLPPPGFG\FA\QNPMVPWPPVV\HLP\GQPQRMMGPLSQASRYIGPQNFYQVKDIRRPERRHSDPWGRQDQQQLDRPFNRGKGDRQRFYSDSHHLKRERHEKEWEQESERHRRRDRSQDKDRDRKSREEGHKDKERARLSHGDRGTDGKASRDSRNVDKKPDKPKSEDYEKDKEREKSKHREGEKDRDRYHKDRDHTDRTKSKR
3539	A	157	1769	GSWTVELSLKPSASPSLKWVCLPGAAAVNKHRSGAGGLIRSLIQCTWAPAGPARRGGRGIEDFPYLFFQLTHCQQRICSVTQAGVQWCDHSSLQPQTPGLNQSSHLSLLSSRDYRMLSSFNEWFWQDRFWLPPNVTWTELEDRDGRVYPHPQDLLAALPLALVLLAMRLAFERFIGLPLSRWLGVRDQTRRQVKPNATLEKHFLTEGHRPKEPQLSLLAAQCGLTLQQTQRWFRRRRNQDRPQLTKKFCEASWRFLFYLSSFVGGLSVLYHESWLWAPVMCWDRYPNQLTLSCPAADSEA\SLYWWYLLELGFYLSLLIRLPFDVKRKGGGPSSIKPRPHYDPPSTA\DFKEQVIHHFVAVILMTFSYSANLLRIGSLVLLLHDSSDYLLEACKMVNYMQYQQVCDALFLIFSFVFFYTRLVLFPTQILYTTYYESISNRGPFFGYYFFNGLLMLLQLLHVFWSCLILRMLYSFMKKGQMEKDIRSDVEESDSSEEAAAAQEPLQLKNGTAGGPRPAPTDGPRSRVAGRLTNRHTTAT
3540	A	267	1397	SPAGYCHSGLLPGCSRSA/CADLAKHQELPGKKLLSEKKLKRYFVDYRRVLVCGGNGGAGASCFHSEPRKEFGGPDGGDGGNGGHVILRVDQQVKSLSSVLSRYQGFSGEDGGSKNCFGRSGAVLYIRVPVGTLVKEGGRVVADLSCVGDEYIAALGGAGGKGNRFFLANNNRAPVTCTPGQPGQQRVLHLELKTVAHAGMVGFPNAGKSSLLRAISNARPAVASYPFTTLKPHVGIVHYEGHLQIAVADIPGIIRGAHQNRGLGSAFLRHIERCRFLLFVVDLSQPEPWTQVDDLKYELEMYEKGLSARPHAIVANKIDLPEAQANLSQLRDHLGQEVIVLSALTGENLEQLLLHLKVLYDAYAEAELGQGRQPLRW
3541	A	1	8008	DTQVSETLKRFAGKVTTASVKERREILSELGKCVAGKDLPEGAVKGLCKLFCLTLHRYRDAASRRALQAAIQQLAEAQPEATAKNLLHSLQSSGIGSKAGVPSKSSGSAALLALTWTCLLVRIVFPSRAKRQGDIWNKLVEVQCLLLLEVLGGSHKHAVDGAVKKLT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KLWKENPGLVEQYLSAILSLEPNQNYAGMLGLL VQFCTSHKEMDVVSQHKSALLDFYMKNILMSK VKPPKYLLDSCAPLLRYLSHSEFKDLILPTIQKSL LRSPENVIETISSLLASVTLDLSQYAMDIVKGLAG HLKSNSPRLMDEAVLALRNLARQCSDSSAMESL TKHLFAILGGSEGKLTVVAQKMSVLSGIGSVSHH VVSGPSSQVLNGIVAELFIPFLQQEVHEGTLVHA VSVLALWCNRFTMEVPKKLTEWFKKAFSLKTST SAVRHAYLQCMLASYRGDTLLQALDLLPLLIQT VEKAASQSTQVPTITEGVAAALLLLKLSVADSQA EAKLSSFWQLIVDEKKQVFTSEKFLVMASEDAL CTVLH\LTERLFLDHPHRLTGNKVQQYHRALVA VLLSRTWHVRRQAQQTVRKLLSSLGGFKLAHGL LEELKTVLSSHKVLPLEALVTDAGEVTEAGKAY VPPRVLQEALCVISGVPGLKGDVTDTEQLAQEM LIISHHPSLVAVQSGLWPALLARMKIDPEAFITRH LDQIIPRMTTQSPLNQSSMNAMGSLSVLSPDRVL PQLISTITASVQNPALRLVTREEFAIMQTPAGELY DKSIIQSAQQDSIKKANMKRENKAYSFKEQIIELE LKEEIKKKKGIKEEVQLTSKQKEMLQAQLDREA QVRRRLQELDGELEAALGLLDIILAKNPSGLTQYI PVLVDSFLPLLKSPLAAPRIKNPFLSLAACVMPSR LKALGTLVSHVTLRLLKPECVLDKSWCQEELSV AVKRAVMLLHTHTITSRVGKGEPGAAPLSAPAFS LVFPFLKMVLTEMPHHSEEEEEWMAQILQILTVQ AQLRASPNTPPGRVDENGPELLPRVAMLRLLTW VIGTGSPRLQVLASDTLTTLCASSSGDDGCAFAE QEEVDVLLCALQSPCASVRETVLRGLMELHMVL PAPDTDEKNGLNLLRRLWVVKFDKEEEIRKLAE RLWSMMGLDLQPDLCSLLIDDVIYHEAAVRQAG AEALSQAVARYQRQAAEVMGRLMEIYQEKLYR PPPVLDALGRVISESPPDQWEARCGLALALNKLS QYLDSSQVKPLFQFFVPDALNDRHPDVRKCMLD AALATLNTHGKENVNSLLPVFEEFLKNAPNDAS YDAVRQSVVVLMGSLAKHLDKSDPKVKPIVAKL IAALSTPSQQVQESVASCLPPLVPAIKEDAGGMIQ RLMQQLLESDKYAERKGAAYGLAGLVKGLGILS LKQQEMMAALTDAIQDKKNFRRREGALFAFEM LCTMLGKLFEPYVVHVLPHLLLCFGDGNQYVRE AADDCAKAVMSNLSAHGVKLVLPSLLAALEEES WRTKAGSVELLGAMAYCAPKQLSSCLPNIVPKL TEVLTDSHVKVQKAGQQALRQIGSVIRNPEILAI APVLLDALTDPSRKTQKCLQTLLDTKFVHFIDAP SLALIMPIVQRAFQDRSTDTRKMAAQIIGNMYSL TDQKDLAPYLPSVTPGLKASLLDPVPEVRTVSAK ALGAMVKGMGESCFEDLLPWLMETLTYEQSSV DRSGAAQGLAEVMAGLGVEKLEKLMPEIVATAS KVDIAPHVRDGYIMMFNYLPITFGDKFTPYVGPII PCILKALADENEFVRDTALRAGQRVISMYAETAI ALLLPQLEQGLFDDLWRIRFSSVQLLGDLLFHISG VTGKMTTETASEDDNFGTAQSNKAIITALGVERR NRVLAGLYMGRSDTQLVVRQASLHVWKIVVSN TPRTLREILPTLFGLLLGFLASTCADKRTIAARTL GDLVRKLGEKILPEIIPILEEGLRSQKSDERQGVCI GLSEIMKSTSRDAVLYFSESLVPTARKALCDPLE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EVREAAAKTFEQLHSTIGHQALEDILPFLLKQLD DEEVSEFALDGLKQVMAIKSRVVLPYLVPKLTTP PVNTRVLAFLSSVAGDALTRHLGVILPAVMLAL KEKLGTPDEQLEMANCQAVILSVEDDTGHRIIIE DLLEATRSPEVGMRQAAAIILNIYCSRSKADYTS HLRSLVSGLIRLFNDSSPVVLEESWDALNAITKK LDAGNQLALIEELHKEIRLIGNESKGEHVPGFCLP KKGVTSILPVLREGVLTGSPEQKEEAAKALGLVI RLTSADALRPSVVSITGPLIRILGDRFSWNVKAAL LETLSLLLAKVGIALKPFLPQLQTTFTKALQDSNR GVRLKAADALGKLISIHIKVDPLFTELLNGIRAME DPGVRDTMLQALRFVIQGAGAKVDAVIRKNIVS LLLSMLGHDEDNTRISSAGCLGELCAFLTEEELS AVLQQCLLADVSGIDWMVRHGRSLALSVAVNV APGRLCAGRYSSDVQEMILSSATADRIPIAVSGV RGMGFLMRHHIETGGGQLPAKLSSLFVKCLQNP SSDIRLVAEKMIWWANKDPLPPLDPQAIKPILKA LLDNTKDKNTVVRAYSDQAIVNLLKMRQGEEVF QSLSKILDVASLEVLNEVNRRSLKKLASQADSTE QVDDTILT
3542	A	62	1130	PWNPQDFPGNRGLMG\QKGEIGPP\GQQGKKGAP GMP\GLMGSNGSPGQPGTPGSKGSKGEPGIQGMP GASGLKGEPGATGSPGEPGYMGLPGIQGKKGDK GNQGEKGIQGQKGENGRQGIPGQQGIQGHHGAK GERGEKGEPGVRGAIGSKGESGVDGLMGPAGPK GQPGDPGPQGPPGLDGKPGREFSEQFIRQVCTDV IRAQLPVLLQSGRIRNCDHCLSQHGSPGIPGPPGPI GPEGPRGLPGLPGRDGVPGLVGVPGRPGVRGLK GLPGRNGEKGSQGFGYPGEQGPPGPPGPEGPPGI SKEGPPGDPGLPGKDGDHGKPGIQGQPGPPGICD PSLCFSVIARRDPFRKGPNY
3543	A	654	194	PARSLEKMKASVVLSLLGYLVVPSGAYILGRCTV AKKLHDGGLDYFERYSLENWVCLAYFESKFNPS\ AIYENTREGYTGFGLFQMRGSDWCGDHGRNRC HMSCSALLNPNLEKTIKCAKTIVKGKEGMGAWP TWSRYCQYSDTLARWLDGCKL
3544	A	2	1074	SCRLAAGRLAQWLLRASRSGMLRAGWLRGAAA LALLLAARVVAAFEPITVGLAIGAASAITGYLSY NDIYCRFAECCREERPLNASALKLDLEEKLFGQH LATEVI\FKALTGFRNNKNPKKPLTLSLHGWAGT GKNFVSQMGAENLHPKGLKSNFVHLFVSTLHFP HEQKIKLYQDQLQKWIRGNVSACANSVFIFDEM DKL\HPGIIE\AIKPFLDYYEHVERVSYR\KAIFIFLS NAGGDLITKTALDFWRAGRKREDIQLKDLEPVL SVGVFNNKHSGLWHSGLIDKNLIDYFIPFLPLEYR HVKMCVRAEMRARGSAIDEDIVTRVAEEMTFFP\ RDEKIYSDKGCKTVQSRLDFH
3545	A	3	273	SAQGRSWGRFYRQIKRHPGIIPMIGLICLGMGSA ALYLLRLALRSPDVW*SWDRKNNPEPWNRLSPN DQYKFLAVSTDYKKLKKDRPDF
3546	A	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAVPA PKVPIKMQVKHWPSEQDPEKAWGARVVEPPEK DDQLVVLFPVQKPKLLTTEEKPRGQGRGPILPGT KAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQG EERPRLWVMPNHQVLLGPEEDQDHIYHPQ*GSR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GHHCPRPVPRPRLLGLGPSLPCPS
3547	A	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAVPA PKVPIKMQVKHWPSEQDPEKAWGARVVEPPEK DDQLVVLFPVQKPKLLTTEEKPRGQGRGPILPGT KAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQG EERPRLWVMPNHQVLLGPEEDQDHIYHPQ*GSR GHHCPRPVPRPRLLGLGPSLPCPS
3548	A	3	1641	TWLPSVPAEEVQQPEMAAVLNAERLEVSVDGLT LSPDPEERPGAEGAPLAAATAATALATWIRSRPG RLRGTARSPGRRAAGGAAEEARRLEQRWGFGLE ELYGLALRFFKEKDGKAFHPTYEEKLKLVALHK QVLMGPYNPDTCPEVGFFDVLGNDRRREWAAL GNMSKEDAMVEFVKLLNRCCHLFSTYVASHKIE KEEQEKKRKEEEERRRREEEERERLQKEEEKRRR EEEERLRREEEERRRIEEERLRLEQQKQQIMAAL NSQTAVQFQQYAAQQYPGNYEQQQILIRQLQEQ HYQQYMQQLYQVQLAQQQAALQKQQEVVVAG SSLPTSSKVECNCTQVI*CQFNRQAKTHTDSSEKE LEPEAAEEALENGPKESLPVIAAPSMWTRPQIKD FKEKIQQDADSVITVGRGEVVTVRVPTHEEGSYL FWEFATDNYDIGFGVYFEWTDSPNTAVSVHVSE SSDDDEEEEENIGCEEKAKKNANKPLLDEIVPVY RRDCHEEVYAGSHQYPGRGVYLLKFDNSYSLW RSKSVYYRVYYTR
3549	A	1837	3593	PAVLVLEPASQSRKQQNTASATAQHWSAQIHKE SFLAPVFTKDEQKHRRPYEFEVERDAKARGLEQF SATHGHTPIILNGWHGESAMDLSCSSEGSPGATS PFPVSASTPKIGAISSLQGALGMDLSGILQAGLIHP VTGQIVNGSLRRDDAATRRRRGRRKHVEGGMD LIFLKEQTLQAGILEVHEDPGQATLSTTHPEGPGP ATSAPEPATAASSQAEKSIPSKSLLDWLRQQADY SLEVPGFGANFSDKPKQRRPRCKEPGKLDVSSLS GEERVPAIPKEPGLRGFLPENKFNHTLAEPILRDT GPRRRGRRPRSELLKAPSIVADSPSGMGPLFMNG LIAGMDLVGLQNMRNMPGIPLTGLVGFPAGFAT MPTGEEVKSTLSMLPMMLPGMAAVPQMFGVGG LLSPPMATTCTSTAPASLSSTTKSGTAVTEKTAE DKPSSHDVKTDTLAEDKPGPGPFSDQSEPAITTSS PVAFNPFLIPGVSPGLIYPSMFLSPGMGMALPAM QQARHSEIVGLESQKRKKKKTKGDNPNSHPEPA PSCEREPSGDENCAEPSAPLPAEREHGAQAGEGA LKDSNNDTN
3550	A	287	39	QLNLNKIATSQKHRDFVAESVGEKPVGSLAGIGE VMDKKLEEGCFDKAYVVLGQFLVLKKDEDLF*E WLRDTGGARTRGSRE
3551	A	21	3925	GDLLEVGLPPGLEFPRGICLRGLRRTMSLDFGSV ALPVQNEDEEYDEEDYEREKELQQLLTDLPHDM LDDDLSSPELQYSDCSEDGTDGQPHHPEQLEMS WNEQMLPKSQSVNGPSCQGLEPYNKVTYKPYQS SAQNNGSPAQEITGSDTFEGLQQQFLGANENSAE NMQIIQLQVLNKAKERQLENLIEKLNESERQIRY LNHQLVIIKDEKDGLTLSLRESQKLFQNGKEREIQ LEAQIKALETQIQALKVNEEQMIKKSRTTEMALE SLKQQLVDLHHSESLQRAREQHESIVMGLTKKY EEQVLSLQKNLDATVTALKEQEDICSRLKDHVK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QLERNQEAIKLEKTEIINKLTRSLEESQKQCAHLL QSGSVQEVAQLQFQLQQAQKAHAMSANMNKA LQEELTELKDEISLYESAAKLGIHPSDSEGELNIEL TESYVDLGIKKVNWKKSKVTSIVQEEDPNEELSK DEFILKLKAEVQRLLGSNSMKRHLVSQLQNDLK DCHKKIEDLHQVKKDEKSIEVETKTDTSEKPKNQ LWPESSTSDVVRDDILLLKNEIQVLQQQNQELKE TEGKLRNTNQDLCNQMRQMVQDFDHDKQEAV DRCERTYQQHHEAMKTQIRESLLAKHALEKQQL FEAYERTHLQLRSELDKLNKEVTAVQECYLEVC REKDNLELTLRKTTEKEQQTQEKIKEKLIQQLEK EWQSKLDQTIKAMKKKTLDCGSQTDQVTTSDVI SKKEMAIMIEEQKCTIQQNLEQEKDIAIKGAMKK LEIELELKHCENITKQVEIAVQNAHQRWLGELPE LAEYQALVKAEQKKWEEQHEVSVNKRISFAVSE AKEKWKSELENMRKNILPGKELEEKIHSLQKELE LKNEEVPVVIRAELAKARSEWNKEKQEEIHRIQE QNEQDYRQFLDDHRNKINEVLAAAKEDFMKQK TELLLQKETELQTCLDQSRREWTMQEAKRIQLEI YQYEEDILTVLGVLLSDTQKEHISDSEDKQLLEI MSTCSSKWMSVQYFEKLKGCIQKAFQDTLPLLV ENADPEWKKRNMAELSKDSASQGTGQGDPGPA AGHHAQPLALQATEAEADKKKVLEIKDLCCGHC FQELEKAKQECQDLKGKLEKCCRHLQHLERKHK AVVEKIGEENNKVVEELIEENNDMKNKLEELQT LCKTPPRSLSAGAIENACLPCSGGALEELRGQYIK AVKKIKCDMLRYIQESKERAAEMVKAEVL*ERQ ETARKMRKYYLICLQQILQDDGKEGAEKKIMNA ASKLATMAKLLETPISSKSQSKTTQSGMSK
3552	A	771	375	ARTRQTSGQAREPEKESPAPGGGGLAEIRSRQQL SQTSRIPPLAKDQAVEAMFPPARGKELLSFEDVA MYFTREEWGHLNWGQKDLYRDVMLENYRNMV LLVYFQFDAAIPLC*TSLAHSSWLQLYFRLYF
3553	A	76	72	PGVRGVEAPGGVAPGRNAMRRGERRDAGGPRP ESPVPAGRASLEEPPDGPSAGQATGPGEGRRSTE SEVYDDGTNTFFWRAHTLTVLFILTCTLGYVTLL EETPQDTAYNTKRGIVASILVFLCFGVTQAKDGP FSRPHPAYWRFWLCVSVVYELFLIFILFQTVQDG RQFLKYVDPKLGVPLPERDYGGNCLIYDPDNET DPFHNIWDKLDGFVPAHFLGWYLKTLMIRDWW MCMIISVMFEFLEYSLEHQLPNFSECWWDHWIM DVLVCNGLGIYCGMKTLEWLSLKTYKWQGLWN IPTYKGKMKRIAFQFTPYSWVRFEWKPASSLRR WLAVCGIILVFLLAELNTFYLKFVLWMPPEHYLV LLRLVFFVNVGGVAMREIYDFMDDPKPHKKLGP QAWLVAAITATELLIVVKYDPHTLTLSLPFYISQC WTLGSVLALTWTVWRFFLRDITLRYKETRWQK WQNKDDQGSTVGNGDQHPLGLDEDLLGPGVAE GEGAPTPN*PRGPAPRPLPSAPRAVCGASSRR
3554	A	2	2106	FDEFSALPSPSLQTSWSFGPMSRRALRRLRGEQR GQEPLGPGALHFDLRDDDDAEEEGPKRELGVRR PGGAGKEGVRVNNRFELINIDDLEDDPVVNGERS GCALTDAVAPGNKGRGQRGNTESKTDGDDTET VPSEQSHASGKLRKKKKKQKNKKSSTGEASENG LEDIDRILERIEDSTGLNRPGPAPLSSRKHVLYVE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HRHLNPDTELKRYFGARAILGEQRPRQRQRVYP KCTWLTTPKSTWPRYSKPGLSMRLLESKKGLSFF AFEHSEEYQQAQHKFLVAVESMEPNNIVVLLQT SPYHVDSLLQLSDACRFQEDQEMARDLVERALY SMECAFHPLFSLTSGACRLDYRRPENRSFYLALY KQMSFLEKRGCPRTALEYCKLILSLEPDEDPLCM LLLIDHLALRARNYEYLIRLFQEWEVGASLAHRN LSQLPNFAFSVPLAYFLLSQQTDLPECEQSSARQ KASLLIQQALTMFPGVLLPLLESCSVRPDASVSSH RFFGPNAEISQPPALSQLVNLYLGRSHFLWKEPA TMSWLEENVHEVLQAVDAGDPAVEACENRRKV LYQRAPRNIHRHVILSEIKEAVAALPPDVTTQSV MGFDPLPPSDTIYSYVRPERLSPISHGNTIALFFRS LLPNYTMEGERPEEGVAGGLNRNQGLNRLMLA VRDMMANFHLNDLEAPHEDDA*GEGEWD
3555	A	2	2106	FDEFSALPSPSLQTSWSFGPMSRRALRRLRGEQR GQEPLGPGALHFDLRDDDDAEEEGPKRELGVRR PGGAGKEGVRVNNRFELINIDDLEDDPVVNGERS GCALTDAVAPGNKGRGQRGNTESKTDGDDTET VPSEQSHASGKLRKKKKKQKNKKSSTGEASENG LEDIDRILERIEDSTGLNRPGPAPLSSRKHVLYVE HRHLNPDTELKRYFGARAILGEQRPRQRQRVYP KCTWLTTPKSTWPRYSKPGLSMRLLESKKGLSFF AFEHSEEYQQAQHKFLVAVESMEPNNIVVLLQT SPYHVDSLLQLSDACRFQEDQEMARDLVERALY SMECAFHPLFSLTSGACRLDYRRPENRSFYLALY KQMSFLEKRGCPRTALEYCKLILSLEPDEDPLCM LLLIDHLALRARNYEYLIRLFQEWEVGASLAHRN LSQLPNFAFSVPLAYFLLSQQTDLPECEQSSARQ KASLLIQQALTMFPGVLLPLLESCSVRPDASVSSH RFFGPNAEISQPPALSQLVNLYLGRSHFLWKEPA TMSWLEENVHEVLQAVDAGDPAVEACENRRKV LYQRAPRNIHRHVILSEIKEAVAALPPDVTTQSV MGFDPLPPSDTIYSYVRPERLSPISHGNTIALFFRS LLPNYTMEGERPEEGVAGGLNRNQGLNRLMLA VRDMMANFHLNDLEAPHEDDA*GEGEWD
3556	A	3388	1650	KTRGTMFYYPNVLQRHTGCFATIWLAATRGSRL VKREYLRVNVVKTCEEILNYVLVRVQPPQPGLP RPRFSLYLSAQLQIGVIRVYSQQCQYLVEDIQHIL ERLHRAQLQIRIDMETELPSLLLPNHLAMMETLE DAPDPFFGMMSVDPRLPSPFDIPQIRHLLEAAIPE RVEEIPPEVPTEPREPERIPVTVLPPEAITILEAEPIR MLEIEGERELPEVSRRELDLLIAEEEEAILLEIPRL PPPAPAE*GQELLDQVGCQCWEGSPHFSCPFPLR VEGMGEALGPEELRLTGWEPGALLMEVTPPEEL RLPAPPSPERRPPVPPPPRRRRRRRLLFWDKETQI SPEKFQEQLQTRAHCWECPMVQPPERTIRGPAEL FRTPTLSGWLPPELLGLWTHCAQPPPKALRRELP EEAAAEEERRKIEVPSEIEVPREALEPSVPLMVSL EISLEAAEEEKSRISLIPPEERWAWPEVEAPEAPA LPVVPELPEVPMEMPLVLPPELELLSLEAVHRAV ALELQANREPDFSSLVSPLSPRRMAARVFYLLLV LSAQQILHVKQEKPYGRLLIQPGPRFH
3557	A	3388	1650	KTRGTMFYYPNVLQRHTGCFATIWLAATRGSRL VKREYLRVNVVKTCEEILNYVLVRVQPPQPGLP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RPRFSLYLSAQLQIGVIRVYSQQCQYLVEDIQHIL ERLHRAQLQIRIDMETELPSLLLPNHLAMMETLE DAPDPFFGMMSVDPRLPSPFDIPQIRHLLEAAIPE RVEEIPPEVPTEPREPERIPVTVLPPEAITILEAEPIR MLEIEGERELPEVSRRELDLLIAEEEEAILLEIPRL PPPAPAE*GQELLDQVGCQCWEGSPHFSCPFPLR VEGMGEALGPEELRLTGWEPGALLMEVTPPEEL RLPAPPSPERRPPVPPPPRRRRRRRLLFWDKETQI SPEKFQEQLQTRAHCWECPMVQPPERTIRGPAEL FRTPTLSGWLPPELLGLWTHCAQPPPKALRRELP EEAAAEEERRKIEVPSEIEVPREALEPSVPLMVSL EISLEAAEEEKSRISLIPPEERWAWPEVEAPEAPA LPVVPELPEVPMEMPLVLPPELELLSLEAVHRAV ALELQANREPDFSSLVSPLSPRRMAARVFYLLLV LSAQQILHVKQEKPYGRLLIQPGPRFH
3558	A	489	2360	IRPRPRGRRRALDSPNAAAPPVYVCRSPGEPTSL VNMASEDIAKLAETLAKTQVAGGQLSFKGKSLK LNTAEDAKDVIKEIEDFDSLEALRLEGNTVGVEA ARVIAKAL*KKSELKRCHWSDMFTGRLRTEIPPA LISLGEGLITAGAQLVELDLSDNAFGPDGVQGFE ALLKSSACFTLQELKLNNCGMGIGGGKILAAALT ECHRKSSAQGKPLALKVFVAGRNRLENDGATAL AEAFRVIGTLEEVHMPQNGINHPGITALAQAFAV NPLLRVINLNDNTFTEKGAVAMAETLKTLRQVE VINFGDCLVRSKGAVAIADAIRGGLPKLKELNLS FCEIKRDAALAVAEAMADKAELEKLDLNGNTLG EEGCEQLQEVLEGFNMAKVLASLSDDEDEEEEE EGEEEEEEAEEEEEEDEEEEEEEEEEEEEEPQQRG QGEKSATPSRKILDPNTGEPAPVLSSPPPADVSTF LAFPSPEKLLRLGPKSSVLIAQQTDTSDPEKVVSA FLKVSSVFKDEATVRMAVQDAVDALMQKAFNS SSFNSNTFLTRLLVHMGLLKSEDKVKAIANLYGP LMALNHMVQQDYFPKALAPLLLAFVTKPNSALE SCSFARHSLLQTLYKV
3559	A	489	2360	IRPRPRGRRRALDSPNAAAPPVYVCRSPGEPTSL VNMASEDIAKLAETLAKTQVAGGQLSFKGKSLK LNTAEDAKDVIKEIEDFDSLEALRLEGNTVGVEA ARVIAKAL*KKSELKRCHWSDMFTGRLRTEIPPA LISLGEGLITAGAQLVELDLSDNAFGPDGVQGFE ALLKSSACFTLQELKLNNCGMGIGGGKILAAALT ECHRKSSAQGKPLALKVFVAGRNRLENDGATAL AEAFRVIGTLEEVHMPQNGINHPGITALAQAFAV NPLLRVINLNDNTFTEKGAVAMAETLKTLRQVE VINFGDCLVRSKGAVAIADAIRGGLPKLKELNLS FCEIKRDAALAVAEAMADKAELEKLDLNGNTLG EEGCEQLQEVLEGFNMAKVLASLSDDEDEEEEE EGEEEEEEAEEEEEEDEEEEEEEEEEEEEEPQQRG QGEKSATPSRKILDPNTGEPAPVLSSPPPADVSTF LAFPSPEKLLRLGPKSSVLIAQQTDTSDPEKVVSA FLKVSSVFKDEATVRMAVQDAVDALMQKAFNS SSFNSNTFLTRLLVHMGLLKSEDKVKAIANLYGP LMALNHMVQQDYFPKALAPLLLAFVTKPNSALE SCSFARHSLLQTLYKV
3560	A	2	1198	FVRELPRPRPGAATAAIMVSVINTVDTSHEDMIH DAQMDYYGTRLATCSSDRSVKIFDVRNGGQILIA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DLRGHEGPVWQVAWAHPMYGNILASCSYDRKV IIWREENGTWEKSHEHAGHDSSVNSVCWAPHDY GLILACGSSDGAISLLTYTGEGQWEVKKINNAHT IGCNAVSWAPAVVPGSLIDHPSGQKPNYIKRFAS GGCDNLIKLWKEEEDGQWKEEQKLEAHSDWVR DVAWAPSIGLPTSTIASCSQDGRVFIWTCDDASS NTWSPKLLHKFNDVVWHVSWSITANILAVSGGD NKVTLWKESVDGQWVCISDVNKGQGSVSASVT EGQQNEQ*QDRWGLAPHPPAPGLPLPGPTNQTT GKSPQLQQDYFPRRSYRCSHRLIICLNVIGDAL
3561	A	540	86	WRVKEMTSTLPKALGRKTASRSHTTLQGGSCCP VLWTAKLRCRKLRFPLPPPPPSSSAWPWQGWGI RGEQEAEGPLGETGPPVGPELSGLRQWRKLIKGR YGEWRGSGQKTGQPS*TTMQGGETEENRTETTT GNKQRESEAPWVRHTYIT
3562	A	1920	242	PMMAMPFFERFKSSIQRPSPVLVLSQNTKRESGR KVQSGNINAAKTIADIIRTCLGPKSMMKMLLDP MGGIVMTNDGNAILREIQVQHPAAKSMIEISRTQ DEEVGDGTTSVIILAGEMLSVAEHFLEQQMHPTV VISAYRKALDDMISTLKKISIPVDISDSDMMLNIIN SSITTKAISRWSSLACNIALDAVKMVQFEENGRK EIDIKKYARVEKIPGGIIEDSCVLRGVMINKDVTH PRMRRYIKNPRIVLLDSSLEYKKGESQTDIEITRE EDFTRILQMEEEYIQQLCEDIIQLKPDVVITEKGIS DLAQHYLMRANITAIRRVRKTDNNRIARACGARI VSRPEELREDDVGTGAGLLEIKKIGDEYFTFITDC KDPKACTILLRGASKEILSEVERNFQDAMQVCRN VLLDPQLVPGGGASEMAVAHALTEKSKAMTGV EQWPYRAVAQALEVIPRTLIQNCGASTIRLLTSLR AKHTQENCETWGVNGETGTLVDMKELGIWEPL AVKLQTYKTAVETAVLLLRIDDIVSGHKKKGDD QSRQGGAPDAGQE
3563	A	1571	560	GPSLLGTRGTPNPARTLQIFFLIIGRRLTGRMAAV DDLQFEEFGNAATSLTANPDATTVNIEDPGETPK HQPGSPRGSGREEDDELLGNDDSDKTELLAGQK KSSPFWTFEYYQTFFDVDTYQVFDRIKGSLLPIPG KNFVRLYIRSNPDLYGPFWICATLVFAIAISGNLS NFLIHLGEKTYHYVPEFRKVSIAATIIYAYAWLVP LALWGFLMWRNSKVMNIVSYSFLEIVCVYGYSL FIYIPTAILWIIPHKAVRWILVMIALGISGSLLAMT FWPAVREDNRRVALATIVTIVLLHMLLSVGCLA YFFDAPEMDHLPTTTATPNQTVAAAKSS
3564	A	1	328	NSRVDDFVAHLQRPLLGPASCLGILRPAMTAHSF ALPGIIFTTFWGLVGIAGPWFVPKGPNRGVIITML VATAVCCYLFWLIAILAQLNPLFGPQLKNETIWY VRFLWE
3565	A	2	1081	FVTDFPARSMAATSLMSALAARLLQPAHSCSLRL RPFHLAAVRNEAVVISGRKLAQQIKQEVRQEVEE WVASGNKRPHLSVILVGENPASHSYVLNKTRAA AVVGINSETIMKPASISEEELLNLINKLNNDDNVD GLLVQLPLPEHIDERRICNAVSPDKDVDGFHVIN VGRMCLDQYSMLPATPWGVWEIIKRTGIPTLGK NVVVAGRSKNVGMPIAMLLHTDGAHERPGGDA TVTISHRYTPKEQLKKHTILADIVISAAGIPNLITA DMIKEGAAVIDVGINRVHDPVTAKPKLVGDVDF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EGVRQKAGYITPVPGGVGPMTVAMLMKNTIIAA KKVLRLEEREVLKSKELGVATN
3566	A	3	1130	SCRRGRQQQRRNVSLSSQFAHTMAAPAQQTTQP GGGKRKGKAQYVLAKRARRCDAGGPRQLEPGL QGILITCNMNERKCVEEAYSLLNEYGDDMYGPE KFTDKDQQPSGSEGEDDDAEAALKKEVGDIKAS TEMRLRRFQSVESGANNVVFIRTLGIEPEKLVHHI LQDMYKTKKKKTRVILRMLPISGTCKAFLEDMK KYAETFLEPWFKAPNKGTFQIVYKSRNNSHVNR EEVIRELAGIVCTLNSENKVDLTNPQYTVVVEIIK AVCCLSVVKDYMLFRKYNLQEVVKSPKDPSQLN SKQGNGKEAKLESADKSDQNNTAEGKNNQQVP ENTEELGQTKPTSNPQVVNEGGAKPELASQATE GSKSNENDFS
3567	A	248	3498	GKKDSSPWTCPFHPPLQLFFVIRNTRQLGDFHLA KIKVRNYWTADGDLDIGAKNVKLYVNRNLIFNG KLDKGDREAPADHSILVDQKNEKSEQLEEAMNA HSEESKGTHEMAGASGDKELGLGCSPPAETLAD AKLSSQGNVSGKRKNSTNCRKDSLSQLEEYLRLS AVPTSMGDMPSAPATSPPVKCPPVHEEPSLIQQL ENLMGRKICEPPGKTPSWLQPSPTGKDRKQGGR KPKPLWLSPEKPLAWKGRLPSDDVIGEGPGETEA RDKGLRHEPGWGTSRSVNTKERPQRATTKVHSD DSDIFNQPPNRERPASGRRGSRKDAGSSSHGDDQ PASREDTWSSRTPSRSRWRSEQEHTLHESWSSLS AFDRSHRGRISNTELPGDILDELLQQKSSRHSDLP PSKKGEQPGLSRGQDGYSGETDAGGDFKIPVLPY GQRLVIDIKSTWGDRHYVGLNGIEIFSSKGEPVQI SNIKADPPDINILPAYGKDPRVVTNLIDGVNRTQ DDMHVWLAPFTRGRSHSITIDFTHPCHVALIRIW NYNKSRIHSFRGVKDITMLLDTQCIFEGEIAKASG TLAGAPEHFGDTILFTTDDDILEAIFYSDEMFDLD VGSLDSLQDEEAMRRPSTADGEGDERPFTQAGL GADERIPELELPSSSPVPQVTTPEPGIYHGICLQLN FTASWGDLHYLGLTGLEVVGKEGQALPIHLHQIS ASPRDLNELPEYSDDSRTLDKLIDGTNITMEDEH MWLIPFSPGLDHVVTIRLDRAESIAGLRFWNYNK SPEDTYRGAKIVHVSLDGLCVSPPEGFLIRKGPG NCHFDFAQEILFVDYLRAQLLPQPARRLDMRSLE CASMDYEAPLMPCGFIFQFQLLTSWGDPYYIGLT GLELYDERGEKIPLSENNIAAFPDSVNSLEGVGG DVRTPDKLIDQVNDTSDGRHMWLAPILPGLVNR VYVIFDLPTTVSMIKLWNYAKTPHRGVKEFGLL VDDLLVYNGILAMVSHLVGGILPTCEPTVPYHTI LFTEDRDIRHQEKHTTISNQAEDQDVQMMNENQ IITNAKRKQSVVDPALRPKTCISEKETRRRRC
3568	A	50	1724	AQGGTLSAASRFCRGGLLGPWLHPASEMAATLD LKSKEEKDAELDKRIEALRRKNEALIRRYQEIEE DRKKAELEGVAVTAPRKGRSVEKENVAVESEKN LGPSRRSPGTPRPPGASKGGRTPPQQGGRAGMG RASRSWEGSPGEQPRGGGAGGRGRRGRGRGSPH LSGAGDTSISDRKSKEWEERRRQNIEKMNEEME KIAEYERNQREGVLEPNPVRNFLDDPRRRSGPLE ESERDRREESRRHGRNWGGPDFERVRCGLEHER QGRRAGLGSAGDMTLSMTGRERSEYLRWKQER

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EKIDQERLQRHRKPTGQWRREWDAEKTDGMFK DGPVPAHEPSHRYDDQAWARPPKPPTFGEFLSQ HKAEASSRRRRKSSRPQAKAAPRAYSDHDDRWE TKEGAASPAPETPQPTSPETSPKETPMQPPEIPAP AHRPPEDEGEENEGEEDEEWEDISEDEEEEEIEVE EGDEEEPAQDHQAPEAAPTGIPCSEQAHGVPFSP EEPLLEPQAPGTPSSPFSPPSGHQPVSDWGEEVEL NSPRTTHLAGALSPGEAWPFESV
3569	A	1	912	MGRVGRAGVQLGRRRTTWAAERTGQAAAGGP GRALRGQRPDLRSGGAADSPAAGRGELYCGVLP RSPWFLSERRRQMADFDTYDDRAYSSFGGGRGS RGSAGGHGSRSQKELPTEPPYTAYVGNLPFNTV QGDIDAIFKDLSIRSVRLVRDKDTDKFKGFCYVE FDEVDSLKEALTYDGALLGDRSLRVDIAEGRKQ DKGGFGFRKGGPDDRGFRDDFLGGRGGSRPGDR RTGPPMGSRFRDGPPLRGSNMDFREPTEEERAQR PRLQLKPRTVATPLNQVANPNSAIFGGARPREEV VQKEQE
3570	A	1	912	MGRVGRAGVQLGRRRTTWAAERTGQAAAGGP GRALRGQRPDLRSGGAADSPAAGRGELYCGVLP RSPWFLSERRRQMADFDTYDDRAYSSFGGGRGS RGSAGGHGSRSQKELPTEPPYTAYVGNLPFNTV QGDIDAIFKDLSIRSVRLVRDKDTDKFKGFCYVE FDEVDSLKEALTYDGALLGDRSLRVDIAEGRKQ DKGGFGFRKGGPDDRGFRDDFLGGRGGSRPGDR RTGPPMGSRFRDGPPLRGSNMDFREPTEEERAQR PRLQLKPRTVATPLNQVANPNSAIFGGARPREEV VQKEQE
3571	A	28	131	RHFFGNLCAMRAKWRKKRMRRLKRKRRKMRQ RSK
3572	A	3	1202	QSEPHRKVRVDPPVRDRPPPHPPPLLVQRALPGQ GQAEGSDGADGAKRRAMAHQTGIHATEELKEFF AKARAGSVRLIKVVIEDEQLVLGASQEPVGRWD QDYDRAVLPLLDAQQPCYLLYRLDSQNAQGFE WLFLAWSPDNSPVRLKMLYAATRATVKKEFGG GHIKDELFGTVKDDLSFAGYQKHLSSCAAPAPLT SAERELQQIRINEVKTEISVESKHQTLQGLAFPLQ PEAQRALQQLKQKMVNYIQMKLDLERETIELVH TEPTDVAQLPSRVPRDAARYHFFLYKHTHEGDP LESVVFIYSMPGYKCSIKERMLYSSCKSRLLDSV EQDFHLEIAKKIEIGDGAELTAEFLYDEVHPKQH AFKQAFAKPKGPGGKRGHKRLIRGPGENGDDS
3573	A	49	1869	PHCEPNPGAGAMVLLHVLFEHAVGYALLALKEV EEISLLQPQVEESVLNLGKFHSIVRLVAFCPFASS QVALENANAVSEGVVHEDLRLLLETHLPSKKKK VLLGVGDPKIGAAIQEELGYNCQTGGVIAEILRG VRLHFHNLVKGLTDLSACKAQLGLGHSYSRAKV KFNVNRVDNMIIQSISLLDQLDKDINTFSMRVRE WYGYHFPELVKIINDNATYCRLAQFIGNRRELNE DKLEKLEELTMDGAKAKAILDASRSSMGMDISAI DLINIESFSSRVVSLSEYRQSLHTYLRSKMSQVAP SLSALIGEAVGARLIAHAGSLTNLAKYPASTVQIL GAEKALFRALKTRGNTPKYGLIFHSTFIGRAAAK NKGRISRYLANKCSIASRIDCFSEVPTSVFGEKLR EQVEERLSFYETGEIPRKNLDVMKEAMVQAEAE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EAAAEITRKLEKQEKKRLKKEKKRLAALALASS ENSSSTPEECEETSEKPKKKKKQKPQEVPQENGM EDPSISFSKPKKKKSFSKEELMSSDLEETAGSTSIP KRKKSTPKEETVNDPEEAGHRSRSKKKRKFSKEE PVSSGPEEAVGKSSSKKKKKFHKASQED
3574	A	284	2032	CGNERTARLWVQPVVSTMPQASEHRLGRTREPP VNIQPRVGSKLPFAPRARSKERRNPASGPNPMLR PLPPRPGLPDERLKKLELGRGRTSGPRPRGPLRA DHGVPLPGSPPPTVALPLPSRTNLARSKSVSSGDL RPMGIALGGHRGTGELGAALSRLALRPEPPTLRR STSLRRLGGFPGPPTLFSIRTEPPASHGSFHMISAR SSEPFYSDDKMAHHTLLLGSGHVGLRNLGNTCF LNAVLQCLSSTRPLRDFCLRRDFRQEVPGGGRA QELTEAFADVIGALWHPDSCEAVNPTRFRAVFQ KYVPSFSGYSQQDAQEFLKLLMERLHLEINRRGR RAPPILANGPVPSPPRRGGALLEEPELSDDDRANL MWKRYLEREDSKIVDLFVGQLKSCLKCQACGY RSTTFEVFCDLSLPIPKKGFAGGKVSLRDCFNLFT KEEELESENAPVCDRCRQKTRSTKKLTVQRFPRI LVLHLNRFSASRGSIKKSSVGVDFPLQRLSLGDF ASDKAGSPVYQLYALCNHSGSVHYGHYTALCR CQTGWHVYNDSRVSPVSENQVASSEGYVLFYQL MQEPPRCL
3575	A	1	2408	RELDSLADLPERIKPPYANGLSTSHLRSSSVEDVK LIISEGRPTIEVRRCSMPSVICEHTKQFQTISEESN QGSLLTVPGDTSPSPKPEVFSNVPERDLSNVSNIH SSFATSPTGASNSKYVSADRNLIKNTAPVNTVMD SPVHLEPSSQVGVIQNKSWEMPVDRLETLSTRDF ICPNSNIPDQESSLQSFCNSENKVLKENADFLSLR QTELPGNSCAQDPASFMPPQQPCSFPSQSLSDAES ISKHMSLSYVANQEPGILQQKNAVQIISSALDTD NESTKDTENTFVLGDVQKTDAFVPVYSDSTIQEA SPNFEKAYTLPVLPSEKDFNGSDASTQLNTHYAF SKLTYKSSSGHEVENSTTDTQVISHEKENKLESL VLTHLSRCDSDLCEMNAGMPKGNLNEQDPKHC PESEKCLLSIEDEESQQSILSSLENHSQQSTQPEM HKYGQLVKVELEENAEDDKTENQIPQRMTRNK ANTMANQSKQILASCTLLSEKDSESSSPRGRIRLT EDDDPQIHHPRKRKVSRVPQPVQVSPSLLQAKEK TQQSLAAIVDSLKLDEIQPYSSERANPYFEYLHIR KKIEEKRKLLCSVIPQAPQYYDEYVTFNGSYLLD GNPLSKICIPTITPPPSLSDPLKELFRQQEVVRMKL RLQHSIEREKLIVSNEQEVLRVHYRAARTLANQT LPFSACTVLLDAEVYNVPLDSQSDDSKTSVRDRF NARQFMSWLQDVDDKFDKLKTCLLMRQQHEA AALNAVQRLEWQLKLQELDPATYKSISIYEIQEF YVPLVDVNDDFELTPI
3576	A	5	1421	LRLAWHDGARWPLGTPRAAATRREAAALPPVT LALLCLDGVFLSSAENDFVHRIQEELDRFLLQKQ LSKVLLFPPLSSRLRYLIHRTAENFDLLSSFSVGE GWKRRTVICHQDIRVPSSDGLSGPCRAPASCPSR YHGPRPISNQGAAAVPRGARAGRWYRGRKPDQ PLYVPRVLRRQEEWGLTSTSVLKREAPAGRDPEE PGDVGAGDPNSDQGLPVLMTQGTEDLKGPGQR CENEPLLDPVGPEPLGPESQSGKGDMVEMATRF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GSTLQLDLEKGKESLLEKRLVAEEEEDEEEVEED GPSSCSEDDYSELLQEITDNLTKKEIQIEKIHLDTS SFMEELPGEKDLAHVVEIYDFEPALKTEDLLATF SEFQEKGFRIQWVDDTHALGIFPCRASAAEALTR EFSVLKIRPLTQGTKQSKLKALQRPKLLRLVKER PQTNATVARRLVARALGLQHKKKERPAVRGPLP P
3577	A	102	1998	DTRTPGSLEMGPLQFRDVAIEFSLEEWHCLDTAQ RNLYRNVMLENYSNLVFLGIVVSKPDLIAHLEQG KKPLTMKRHEMVANPSGPVICSHFAQDLWPEQN IKDSFQKVILRRYEKRGHGNLQLIKRCESVDECK VHTGGYNGLNQCSTTTQSKVFQCDKYGKVFHK FSNSNRHNIRHTEKKPFKCIECGKAFNQFSTLITH KKIHTGEKPYICEECGKAFKYSSALNTHKRIHTG EKPYKCDKCDKAFIASSTLSKHEIIHTGKKPYKCE ECGKAFNQSSTLTKHKKIHTGEKPYKCEECGKAF NQSSTLTKHKKIHTGEKPYVCEECGKAFKYSRIL TTHKRIHTGEKPYKCNKCGKAFIASSTLSRHEFIH MGKKHYKCEECGKAFIWSSVLTRHKRVHTGEKP YKCEECGKAFKYSSTLSSHKRSHTGEKPYKCEEC GKAFVASSTLSKHEIIHTGKKPYKCEECGKAFNQ SSSLTKHKKIHTGEKPYKCEECGKAFNQSSSLTK HKKIHTGEKPYKCEECGKAFNQSSTLIKHKKIHT REKPYKCEECGKAFHLSTHLTTHKILHTGEKPYR CRECGKAFNHSATLSSHKKIHSGEKPYECDKCG KAFISPSSLSRHEIIHTGEKP
3578	A	1725	445	RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPLRP LSRHPLSSGSPETSAAAIMLLTVRHGTVRYRSSA LLARTKNNIQRYFGTNSVICSKKDKQSVRTEETS KETSESQDSEKENTKKDLLGIIKGMKVELSTVNV RTTKPPKRRPLKSLEATLGRLRRATEYAPKKRIEP LSPELVAAASAVADSLPFDKQTTKSELLSQLQQH EEESRAQRDAKRPKISFSNIISDMKVARSATARV RSRPELRIQFDEGYDNYPGQEKTDDLKKRKNIFT GKRLNIFDMMAVTKEAPETDTSPSLWDVEFAKQ LATVNEQPLQNGFEELIQWTKEGKLWEFPINNEA GFDDDGSEFHEHIFLEKHLESFPKQGPIRHFMELV TCGLSKNPYLSVKQKVEHIEWFRNYFNEKKDILK ESNIQFKLRPWKFLFRNN
3579	A	1725	445	RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPLRP LSRHPLSSGSPETSAAAIMLLTVRHGTVRYRSSA LLARTKNNIQRYFGTNSVICSKKDKQSVRTEETS KETSESQDSEKENTKKDLLGIIKGMKVELSTVNV RTTKPPKRRPLKSLEATLGRLRRATEYAPKKRIEP LSPELVAAASAVADSLPFDKQTTKSELLSQLQQH EEESRAQRDAKRPKISFSNIISDMKVARSATARV RSRPELRIQFDEGYDNYPGQEKTDDLKKRKNIFT GKRLNIFDMMAVTKEAPETDTSPSLWDVEFAKQ LATVNEQPLQNGFEELIQWTKEGKLWEFPINNEA GFDDDGSEFHEHIFLEKHLESFPKQGPIRHFMELV TCGLSKNPYLSVKQKVEHIEWFRNYFNEKKDILK ESNIQFKLRPWKFLFRNN
3580	A	3673	1619	LYCVAPYSRHLLGRMSHLPMKLLRKKIEKRNLK LRQRNLKFQGASNLTLSETQNGDVSEETMGSRK VKKSKQKPMNVGLSETQNGGMSQEAVGNIKVT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KSPQKSTVLTNGEAAMQSSNSESKKKKKKKRK MVNDAEPDTKKAKTENKGKSEEESAETTKETEN NVEKPDNDEDESEVPSLPLGLTGAFEDTSFASLC NLVNENTLKAIKEMGFTNMTEIQHKSIRPLLEGR DLLAAAKTGSGKTLAFLIPAVELIVKLRFMPRNG TGVLILSPTRELAMQTFGVLKELMTHHVHTYGLI MGGSNRSAEAQKLGNGINIIVATPGRLLDHMQN TPGFMYKNLQCLVIDEADRILDVGFEEELKQIIKL LPTRRQTMLFSATQTRKVEDLARISLKKEPLYVG VDDDKANATVDGLEQGYVVCPSEKRFLLLFTFL KKNRKKKLMVFFSSCMSVKYHYELLNYIDLPVL AIHGKQKQNKRTTTFFQFCNADSGTLLCTDVAA RGLDIPEVDWIVQYDPPDDPKEYIHRVGRTARGL NGRGHALLILRPEELGFLRYLKQSKVPLSEFDFS WSKISDIQSQLEKLIEKNYFLHKSAQEAYKSYIRA YDSHSLKQIFNVNNLNLPQVALSFGFKVPPFVDL NVNSNEGKQKKRGGGGGFGYQKTKKVEKSKIF KHISKKSSDSRQFSH
3581	A	23	453	LCRCICIKNITPHCLWDKVLSQFTYILDNLSNFMS HHPHSLRNSCLIRMDLLYWQFTIYTITFCFSHLSG RLTLSAQHISHRPCLLSYSLLFWKVHHLFLEGFPC SPRLDEMSFHQFPQHPVHVSVVHLPIVYKGSMT QVSPH
3582	A	3	950	TRGCGNKMAGKKNVLSSLAVYAEDSEPESDGEA GIEAVGSAAEEKGGLVSDAYGEDDFSRLGGDED GYEEEEDENSRQSEDDDSETEKPEADDPKDNTE AEKRDPQELVASFSERVRNMSPDEIKIPPEPPGRC SNHLQDKIQKLYERKIKEGMDMNYIIQRKKEFRN PSIYEKLIQFCAIDELGTNYPKDMFDPHGWSEDS YYEALAKAQKIEMDKLEKAKKERTKIEFVTGTK KGTTTNATSTTTTTASTAVADAQKRKSKWDSAI PVTTIAQPTILTTTATLPAVVTVTTSASGSKTTVIS AVGTIVKKAKQ
3583	A	3	950	TRGCGNKMAGKKNVLSSLAVYAEDSEPESDGEA GIEAVGSAAEEKGGLVSDAYGEDDFSRLGGDED GYEEEEDENSRQSEDDDSETEKPEADDPKDNTE AEKRDPQELVASFSERVRNMSPDEIKIPPEPPGRC SNHLQDKIQKLYERKIKEGMDMNYIIQRKKEFRN PSIYEKLIQFCAIDELGTNYPKDMFDPHGWSEDS YYEALAKAQKIEMDKLEKAKKERTKIEFVTGTK KGTTTNATSTTTTTASTAVADAQKRKSKWDSAI PVTTIAQPTILTTTATLPAVVTVTTSASGSKTTVIS AVGTIVKKAKQ
3584	A	3	1139	PGSTISSRADRLGAPVLAHPKMAERQEEQRGSPP LRAEGKADAEVKLILYHWTHSFSSQKVRLVIAE KALKCEEHDVSLPLSEHNEPWFMRLNSTGEVPV LIHGENIICEATQIIDYLEQTFLDERTPRLMPDKES MYYPRVQHYRELLDSLPMDAYTHGCILHPELTV DSMIPAYATTRIRSQIGNTESELKKLAEENPDLQE AYIAKQKRLKSKLLDHDNVKYLKKILDELEKVL DQVETELPRRNEETPEEGQQPWLCGESFTLADVS LAVTLHRLKFLGFARRNWGNGKRPNLETYYERV LKRKTFNKVLGHVNNILISAVLPTAFRVAKKRAP KVLGTTLVVGLLAGVGYFAFMLFRKRLGSMILA LRPRPNYF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3585	A	1	1777	RRHSPGSPAFAPSSRATAICPRAARAPATLLLALG AVLWPAAGAWELTILHTNDVHSRLEQTSEDSSK CVNASRCMGGVARLFTKVQQIRRAEPNVLLLDA GDQYQGTIWFTVYKGAEVAHFMNALRYDAMA LGNHEFDNGVEGLIEPLLKEAKFPILSANIKAKGP LASQISGLYLPYKVLPVGDEVVGIVGYTSKETPF LSNPGTNLVFEDEITALQPEVDKLKTLNVNKIIAL GHSGFEMDKLIAQKVRGVDVVVGGHSNTFLYT GNPPSKEVPAGKYPFIVTSDDGRKVPVVQAYAF GKYLGYLKIEFDERGNVISSHGNPILLNSSIPEDPS IKADINKWRIKLDNYSTQELGKTIVYLDGSSQSC RFRECNMGNLICDAMINNNLRHTDEMFWNHVS MCILNGGGIRSPIDERNNGTITWENLAAVLPFGG TFDLVQLKGSTLKKAFEHSVHRYGQSTGEFLQV GGIHVVYDLSRKPGDRVVKLDVLCTKCRVPSYD PLKMDEVYKVILPNFLANGGDGFQMIKDELLRH DSGDQDINVVSTYISKMKVIYPAVEGRIKFSTGS HCHGSFSLIFLSLWAVIFVLYQ
3586	A	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEAQTE MVRTLERKLEAKMIKEESDYHDLESVVQQVEQN LELMTKRAVKAENHVVKLKQEISLLQAQVSNFQ RENEALRCGQGASLTVVKQNADVALQNLRVVM NSAQASIEQLVSGAETLNLVAEILKSIDRISEVKD EEEDS
3587	A	88	1639	GCVGRGLPLPPRHPTPPSSSSSPFVLLAFLLLVRL DPAVSGKMAAPRPPPARLSGVMVPAPIQDLEAL RALTALFKEQRNRETAPRTIFQRVLDILKKSSHA VELACRDPSQVENLASSLQLITECFRCLRNACIEC SVNQNSIRNLDTIGVAVDLILLFRELRVEQESLLT AFRCGLQFLGNIASRNEDSQSIVWVHAFPELFLS CLNHPDKKIVAYSSMILFTSLNHERMKELEENLN IAIDVIDAYQKHPESEWPFLIITDLFLKSPELVQA MFPKLNNQERVTLLDLMIAKITSDEPLTKDDIPVF LRHAELIASTFVDQCKTVLKLASEEPPDDEEALA TIRLLDVLCEMTVNTELLGYLQVFPGLLERVIDL LRVIHVAGKETTNIFSNCGCVRAEGDISNVANGF KSHLIRLIGNLCYKNKDNQDKVNELDGIPLILDN CNISDSNPFLTQWVIYAIRNLTEDNSQNQDLIAK MEEQGLADASLLKKVGFEVEKKGEKLILKSTRD TPKP
3588	A	3	1462	DSPRNRFEILGRPTRTPTRPGPRPAMEDLDALLSD LETTTSHMPRSGAPKERPAEPLTPPPSYGHQPQT GSGESSGASGDKDHLYSTVCKPRSPKPAAPAAPP FSSSSGVLGTGLCELDRLLQELNATQFNITDEIMS QFPSSKVASGEQKEDQSEDKKRPSLPSSPSPGLPK ASATSATLELDRLMASLSDFRVQNHLPASGPTQP PVVSSTNEGSPSPPEPTGKGSLDTMLGLLQSDLSR RGVPTQAKGLCGSCNKPIAGQVVTALGRAWHPE HFVCGGCSTALGGSSFFEKDGAPFCPECYFERFSP RCGFCNQPIRHKMVTALGTHWHPEHFCCVSCGE PFGDEGFHEREGRPYCRRDFLQLFAPRCQGCQGP ILDNYISALSALWHPDCFVCRECFAPFSGGSFFEH EGRPLCENHFHARRGSLCATCGLPVTGRCVSAL GRRFHPDHFTCTFCLRPLTKGSFQERAGKPYCQP CFLKLFG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3589	A	226	6793	SPPKKSRKCNLSFRLISAERWRFFLLILMEMPRKP RLTLFVQRRIENIATEREFDPEEFYYLLEAAEGHA KEGQGIKTDIPRYIISQLGLNKDPLEEMAHLGNY DSGTAETPETDESVSSSNASLKLRRKPRESDFETI KLISNGAYGAVYFVRHKESRQRFAMKKINKQNL ILRNQIQQAFVERDILTFAENPFVVSMYCSFETRR HLCMVMEYVEGGDCATLMKNMGPLPVDMARM YFAETVLALEYLHNYGIVHRDLKPDNLLVTSMG HIKLTDFGLSKVGLMSMTTNLYEGHIEKDAREFL DKQVCGTPEYIAPEVILRQGYGKPVDWWAMGII LYEFLVGCVPFFGDTPEELFGQVISDEINWPEKDE APPPDAQDLITLLLRQNPLERLGTGGAYEVKQHR FFRSLDWNSLLRQKAEFIPQLESEDDTSYFDTRSE KYHHMETEEEDDTNDEDFNVEIRQFSSCSHRFSK VFSSIDRITQNSAEEKEDSVDKTKSTTLPSTETLS WSSEYSEMQQLSTSNSSDTESNRHKLSSGLLPKL AISTEGEQDEAASCPGDPHEEPGKPALPPEECAQ EEPEVTTPASTISSSTLSVGSFSEHLDQINGRSECV DSTDNSSKPSSEPASHMARQRLESTEKKKISGKV TKSLSASALSLMIPGDMFAVSPLGSPMSPHSLSSD PSSSRDSSPSRDSSAASASPHQPIVIHSSGKNYGFT IRAIRVYVGDSDIYTVHHIVWNVEEGSPACQAGL KAGDLITHINGEPVHGLVHTEVIELLLKSGNKVSI TTTPFENTSIKTGPARRNSYKSRMVRRSKKSKKK ESLERRRSLFKKLAKQPSPLLHTSRSFSCLNRSLS SGESLPGSPTHSLSPRSPTPSYRSTPDFPSGTNSSQ SSSPSSSAPNSPAGSGHIRPSTLHGLAPKLGGQRY RSGRRKSAGNIPLSPLARTPSPTPQPTSPQRSPSPL LGHSLGNSKIAQAFPSKMHSPPTIVRHIVRPKSAE PPRSPLLKRVQSEEKLSPSYGSDKKHLCSRKHSL EVTQEEVQREQSQREAPLQSLDENVCDVPPLSRA RPVEQGCLKRPVSRKVGRQESVDDLDRDKLKAK VVVKKADGFPEKQESHQKFHGPGSDLENFALFK LEEREKKVYPKAVERSSTFENKASMQEAPPLGSL LKDALHKQASVRASEGAMSDGPVPAEHRQGGG DFRRAPAPGTLQDGLCHSLDRGISGKGEGTEKSS QAKELLRCEKLDSKLANIDYLRKKMSLEDKEDN LCPVLKPKMTAGSHECLPGNPVRPTGGQQEPPPA SESRAFVSSTHAAQMSAVSFVPLKALTGRVDSGT EKPGLVAPESPVRKSPSEYKLEGRSVSCLEPIEGT LDIALLSGPQASKTELPSPESAQSPSPSGDVRASV PPVLPSSSGKKNDTTSARELSPSSLKMNKSYLLEP WFLPPSRGLQNSPAVSLPDPEFKRDRKGPHPTAR SPGTVMESNPQQREGSSPKHQDHTTDPKLLTCLG QNLHSPDLARPRCPLPPEASPSREKPGLRESSERG PPTARSERSAARADTCREPSMELCFPETAKTSDN SKNLLSVGRTHPDFYTQTQAMEKAWAPGGKTN HKDGPGEARPPPRDNSSLHSAGIPCEKELGKVRR GVEPKPEALLARRSLQPPGIESEKSEKLSSFPSLQ KDGAKEPERKEQPLQRHPSSIPPPPLTAKDLSSPA ARQHCSSPSHASGREPGAKPSTAEPSSSPQDPPKP VAAHSESSSHKPRPGPDPGPPKTKHPDRSLSSQK PSVGATKGKEPATQSLGGSSREGKGHSKSGPDVF PATPGSQNKASDGIGQGEGGPSVPLHTDRAPLDA KPQPTSGGRPLEVLEKPVHLPRPGHPGPSEPADQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KLSAVGEKQTLSPKHPKPSTVKDCPTLCKQTDNRQTDKSPSQPAANTDRRAEGKKCTEALYAPAEGDKLEAGLSFVHSENRLKGAERPAAGVGKGFPEARGKGPGPQKPPTEADKPNGMKRSPSATGQSSFRSTALPEKSLSCSSSFPETRAGVREASAASSDTSSAKAAGGMLELPAPSNRDHRKAQPAGEGRTHMTKSDSLPSFRVSTLPLESHHPDPNTMGGASHRDRALSVTATVGETKGKDPAPAQPPPARKQNVGRDVTKPSPAPNTDRPISLSNEKDFVVRQRRGKESLRSSPHKKAL
3590	A	3	935	RATTRPKNEVQDYVSVEYLSPHMGGTDPFKYSYPPLVDDDFQTPLCENGPITSEDETSSKEDIESDGKETLETISNEEQTPLLKKINPTESTSKAEENEKVDSKVKAFKKPLSVFKGPLLHISPAEELYFGSTESGEKKTLIVLTNVTKNIVAFKVRTTAPEKYRVKPSNSSCDPGASVDIVVSPHGGLTVSAQDRFLIMAAEMEQSSGTGPAELTQFWKEVPRNKVMEHRLRCHTVESSKPNTLTLKDNAFNMSDKTSEDICLQLSRLLESNRKLEDQVQRCIWFQQLLLSLTMLLLAFVTSFFYLLYS
3591	A	303	2	GGSWGPLCPVSPAMSLSDPGLGYHPTCWTLRWPPLCSLHALHVFHCLFSSRLGTPVSPRLAMDPNCSCEAGGSCACAGSCKCKKCKCTSCKKSCCSCCPL
3592	A	1052	1779	GKTMMRKMLLAAALSVTAMTAHADYQCSVTPRDDVIVSPQTVQVKGENGNLVITPDGNVMYNGKQYSLNAAQREQAKDYQAELRSTLPWIDEGAKSRVEKARIALDKIIVQEMGESSKMRSRLTKLDAQVKEQMNRIIETRSDGLTFHYKAIDQVRAEGQQLVNQAMGGILQDSINEMGAKAVLKSGGNPLQNVLGSLGGLQSSIQTEWKKQEKDFQQFGKDVCSRVVTLEDSRKALVGNLK
3593	A	3	1837	LSFEKVDIQTDNDLTKEMYEGKENVSFELQRDFSQETDFSEASLLEKQQEVHSAGNIKKEKSNTIDGTVKDETSPVEECFFSQSSNSYQCHTITGEQPSGCTGLGKSISFDTKLVKHEIINSEERPFKCEELVEPFRCDSQLIQHQENNTEEKPYQCSECGKAFSINEKLIWHQRLHSGEKPFKCVECGKSFSYSSHYITHQTIHSGEKPYQCKMCGKAFSVNGSLSRHQRIHTGEKPYQCKECGNGFSCSSAYITHQRVHTGEKPYECNDCGKAFNGNAKLIQHQRIHTGEKPYECNECGKGFRCSSQLRQHQSIHTGEKPYQCKECGKGFNNNTKLIQHQRIHTASLAEQLFKASGNHPNWGCCLTISSPGPSVYGPKMNMRGAPNSRLAGGREKRTQDTDFGQCSFLPSHSPSCFEPWNVTDYDSSWYRQKQVLSGVWSSPLSILKLPRTLIRISIHIQEMDTPGEMLMTGRGSLGPTLTTEAPAAAQPGKQGPPGTGRCLQAPGTEPGEQTPEGARELSPLQESSSPGGVKAEEEQRAGAEPGTRPSLARSDDNDHEVGALGLQQGKSPGAGNPEPEQDCAARAPVRAEAVRRMPPGAEAGSVVLDD
3594	A	39	261	RAAMMDTSRVQPIKLAIVIKVLGRTGSQGQCTQVRVEFMDDTSRSIIRSVKGPVREGDVLTLLESEREARRLR
3595	A	973	68	GRVGTKHQMADDAGAAGGPGGPGGPGMGNRGGFRGGFGSGIRGRGRGRGRGRGRGRGARGGKAE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DKEWMPVTKLGRLVKDMKIKSLEEIYLFSLPIKE SEIIDFFLGASLKDEVLKIMPVQKQTRAGQRTRF KAFVAIGDYNGHVGLGVKCSKEVATAIRGAIILA KLSIVPVRRGYWGNKIGKPHTVPCKVTGRCGSV LVRLIPAPRGTGIVSAPVPKKLLMMAGIDDCYTS ARGCTATLGNFAKATFDAISKTYSYLTPDLWKE TVFTKSPYQEFTDHLVKTHTRVSVQRTQAPAVA TT
3596	A	106	2960	DERRVGAADMFGRSRSWVGGGHGKTSRNIHSL DHLKYLYHVLTKNTTVTEQNRNLLVETIRSITEIL IWGDQNDSSVFDFFLEKNMFVFFLNILRQKSGRY VCVQLLQTLNILFENISHETSLYYLLSNNYVNSII VHKFDFSDEEIMAYYISFLKTLSLKLNNHTVHFF YNEHTNDFALYTEAIKFFNHPESMVRIAVRTITL NVYKVSLDNQAMLHYIRDKTAVPYFSNLVWFIG SHVIELDDCVQTDEEHRNRGKLSDLVAEHLDHL HYLNDILIINCEFLNDVLTDHLLNRLFLPLYVYSL ENQDKGGERPKISLPVSLYLLSQVFLIIHHAPLVN SLAEVILNGDLSEMYAKTEQDIQRSSAKPSIRCFI KPTETLERSLEMNKHKGKRRVQKRPNYKNVGEE EDEEKGPTEDAQEDAEKAKGTEGGSKGIKTSGES EEIEMVIMERSKLSELAASTSVQEQNTTDEEKSA AATCSESTQWSRPFLDMVYHALDSPDDDYHALF VLCLLYAMSHNKGMDPEKLERIQLPVPNAAEKT TYNHPLAERLIRIMNNAAQPDGKIRLATLELSCL LLKQQVLMSAGCIMKDVHLACLEGAREESVHLV RHFYKGEDIFLDMFEDEYRSMTMKPMNVEYLM MDASILLPPTGTPLTGIDFVKRLPCGDVEKTRRAI RVFFMLRSLSLQLRGEPETQLPLTREEDLIKTDDV LDLNNSDLIACTVITKDGGMVQRSLAVDIYQMS LVEPDVSRLGWGVVKFAGLLQDMQVTGVEDDS RALNITIHKPASSPHSKPFPILQATFIFSDHIRCIIAK QRLAKGRIQARRMKMQRIAALLDLPIQPTTEVLG FGLGSSTSTQHLPFRFYDQGRRGSSDPTVQRSVF ASVDKVPGFAVAQCINEHSSPSLSSQSPPSASGSP SGSGSTSHCDSGGTSSSSTPSTAQSPAGIGHVTQ
3597	A	427	277	GVRRIQHHWAQMHECNVHTYASLFCLFLLHTG KLCCLNSHRHFHCIKYSK
3598	A	1	503	FRPRTKKATAMYLEHYLDSIENLPCELQRNFQL MRELDQRTEDKKAEIDILAAEYISTVKTLSPDQR VERLQKIQNAYSKCKEYSDDKVQLAMQTYEMV DKHIRRLDADLARFEADLKDKMEGSDFESSGGR GLKKGRGQKEKRGSRGRGRRTSEEDTPKKKKH KGG
3599	A	2	3907	KTITALAFSPDGKYLVTGESGHMPAVRVWDVAE HSQVAELQEHKYGVACVAFSPSAKYIVSVGYQH DMIVNVWAWKKNIVVASNKVSSRVTAVSFSED CSYFVTAGNRHIKFWYLDDSKTSKVNATVPLLG RSGLLGELRNNLFTDVACGRGKKADSTFCITSSG LLCEFSDRRLLDKWVELRVYPEVKDSNQACLPP SSFITCSSDNTIRLWNTESSGVHGSTLHRNILSSDL IKIIYVDGNTQALLDTELPGGDKADASLLDPRVGI RSVCVSPNGQHLASGDRMGTLRVHELQSLSEML KVEAHDSEILCLEYSKPDTGLKLLASASRDRLIH VLDAGREYSLQQTLDEHSSSITAVKFAASDGQVR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MISCGADKSIYFRTAQKSGDGVQFTRTHHVVRK TTLYDMDVEPSWKYTAIGCQDRNIRIFNISSGKQ KKLFKGSQGEDGTLIKVQTDPSGIYIATSCSDKNL SIFDFSSGECVATMFGHSEIVTGMKFSNDCKHLIS VSGDSCIFVWRLSSEMTISMRQRLAELRQRQRGG KQQGPSSPQRASGPNRHQAPSMLSPGPALSSDSD KEGEDEGTEEELPALPVLAKSTKKALASVPSPAL PRSLSHWEMSRAQESVGFLDPAPAANPGPRRRG RWVQPGVELSVRSMLDLRQLETLAPSLQDPSQD SLAIIPSGPRKHGQEALETSLTSQNEKPPRPQASQ PCSYPHIIRLLSQEEGVFAQDLEPAPIEDGIVYPEP SDNPTMDTSEFQVQAPARGTLGRVYPGSRSSEK HSPDSACSVDYSSSCLSSPEHPTEDSESTEPLSVD GISSDLEEPAEGDEEEEEEGGMGPYGLQEGSPQ TPDQEQFLKQHFETLASGAAPGAPVQVPERSESR SISSRFLLQVQTRPLREPSPSSSSLALMSRPAQVPQ ASGEQPRGNGANPPGAPPEVEPSSGNPSPQQAAS VLLPRCRLNPDSSWAPKRVATASPFSGLQKAQS VHSLVPQERHEASLQAPSPGALLSREIEAQDGLG SLPPADGRPSRPHSYQNPTTSSMAKISRSISVGEN LGLVAEPQAHAPIRVSPLSKLALPSRAHLVLDIPK PLPDRPTLAAFSPVTKGRAPGEAEKPGFPVGLGK AHSTTERWACLGEGTTPKPRTECQAHPGPSSPCA QQLPVSSLFQGPENLQPPPPEKTPNPMECTKPGA ALSQDSEPAVSLEQCEQLVAELRGSVRQAVRLY HSVAGCKMPSAEQSRIAQLLRDTFSSVRQELEAV AGAVLSSPGSSPGAVGAEQTQALLEQYSELLLRA VERRMERKL
3600	A	1688	916	IPGSTISCSMALCEAAGCGSALLWPRLLLFGDSIT QFSFQQGGWGASLADRLVRKCDVLNRGFSGYN TRWAKIILPRLIRKGNSLDIPVAVTIFFGANDSAL KDENPKQHIPLEEYAANLKSMVQYLKSVDIPENR VILITPTPLCETAWEEQCIIQGCKLNRLNSVVGEY ANACLQVAQDCGTDVLDLWTLMQDSQDFSSYL SDGLHLSPKGNEFLFSHLWPLIEKKVSSLPLLLPY WRDVAEAKPELSLLGDGDH
3601	A	44	223	VHFPLIPQLAKCFWTMNRAARNKSEKRYYSEFL QIAHLFNYGLSSFLREFIIFLIKLLQ
3602	A	37	1124	VPKPASGKRRLEFRPQDSKACAATPHSPGRITSR TRGSQKVRSVPPRLPWAQASASTDWEGLRGVPG PALRRENFLEAAASGRSGRTPTGGVGFRDVGGP HFPIFPAAHFLWCNLHTPRRPACNAPWHSPVGEI SPPPRESQLRRDPEVHFESPAHPLGFRLLPGRGLP ANAVTVETAAMAAPRQIPSHIVRLKPSCSTDSSF TRTPVPTVSLASRELPVSSWQVTEPSSKNLWEQI CKEYEAEQPPFPEGYKVKQEPVITVAPVEEMLFH GFSAEHYFPVSHFTMISRTPCPQDKSETINPKTCS PKEYLETFIFPVLLPGMASLLHQAKKEKCFEVVL QMTPSGGKACVWGHLPSSSHTI
3603	A	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQSVLR RMQKKYWKTKQVFIKATGKKEDEHLVASDAEL DAKLEVFHSVQETCTELLKIIEKYQLRLNGMKS
3604	A	103	2440	QPRRRVFPAAGRGPGRKCSQWGRQASVSFEDVT VDFSKEEWQHLDPAQRRLYWDVTLENYSHLLS VGYQIPKSEAAFKLEQGEGPWMLEGEAPHQSCS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GEAIGKMQQQGIPGGIFFHCERFDQPIGEDSLCSI LEELWQDNDQLEQRQENQNNLLSHVKVLIKERG YEHKNIEKIIHVTTKLVPSIKRLHNCDTILKHTLN SHNHNRNSATKNLGKIFGNGNNFPHSPSSTKNEN AKTGANSCEHDHYEKHLSHKQAPTHHQKIHPEE KLYVCTECVMGFTQKSHLFEHQRIHAGEKSREC DKSNKVFPQKPQVDVHPSVYTGEKPYLCTQCGK VFTLKSNLITHQKIHTGQKPYKCSECGKAFFQRS DLFRHLRIHTGEKPYECSECGKGFSQNSDLSIHQ KTHTGEKHYECNECGKAFTRKSALRMHQRIHTG EKPYVCADCGKAFIQKSHFNTHQRIHTGEKPYEC SDCGKSFTKKSQLHVHQRIHTGEKPYICTECGKV FTHRTNLTTHQKTHTGEKPYMCAECGKAFTDQS NLIKHQKTHTGEKPYKCNGCGKAFIWKSRLKIH QKSHIGERHYECKDCGKAFIQKSTLSVHQRIHTG EKPYVCPECGKAFIQKSHFIAHHRIHTGEKPYECS DCGKCFTKKSQLRVHQKIHTGEKPNICAECGKAF TDRSNLITHQKIHTREKPYECGDCGKTFTWKSRL NIHQKSHTGERHYECSKCGKAFIQKATLSMHQII HTGKKPYACTECQKAFTDRSNLIKHQKMHSGEK RYKASD
3605	A	3	322	SFRMSGRGKGGKGLGKGGAKRHRKVLRDNIQGI TKPAIRRLARRGGVKRISGLIYEETRGVLKVFLEN VIRDAVTYTEHAKRKTVTAMDVVYALKRQGRT LYGFGG
3606	A	1	1749	VPVTAEAKLMGFTQGCVTFEDVAIYFSQEEWGL LDEAQRLLYRDVMLENFALITALVCWHGMEDE ETPEQSVSVEGVPQVRTPEASPSTQKIQSCDMCV PFLTDILHLTDLPGQELYLTGACAVFHQDQKHHS AEKPLESDMDKASFVQCCLFHESGMPFTSSEVG KDFLAPLGILQPQAIANYEKPNKISKCEEAFHVGI SHYKWSQCRRESSHKHTFFHPRVCTGKRLYESS KCGKACCCECSLVQLQRVHPGERPYECSECGKS FSQTSHLNDHRRIHTGERPYVCGQCGKSFSQRAT LIKHHRVHTGERPYECGECGKSFSQSSNLIEHCRI HTGERPYECDECGKAFGSKSTLVRHQRTHTGEK PYECGECGKLFRQSFSLVVHQRIHTTARPYECGQ CGKSFSLKCGLIQHQLIHSGARPFECDECGKSFSQ RTTLNKHHKVHTAERPYVCGECGKAFMFKSKL VRHQRTHTGERPFECSECGKFFRQSYTLVEHQKI HTGLRPYDCGQCGKSFIQKSSLIQHQVVHTGERP YECGKCGKSFTQHSGLILHRKSHTVERPRDSSKC GKPYSPRSNIV
3607	A	92	331	AMAGPGPGPGDPDEQYDFLFKLVLVGDASVGKT CVVQRFKTGAFSERQGSTIGVDFTMKTLEIQGKR VKLQIWDTAGQER
3608	A	545	379	AIKGYIHLSAPRNRYMHTTASNGRMLFMKVTM YMRRGVQIMGWSVRMAFMACFTQ
3609	A	118	873	VWMAWQVSLLELEDRLQCPICLEVFKESLMLQC GHSYCKGCLVSLSYHLDTKVRCPMCWQVVDGS SSLPNVSLAWVIEALRLPGDPEPKVCVHHRNPLS LFCEKDQELICGLCGLLGSHQHHPVTPVSTVCSR MKEELAALFSELKQEQKKVDELIAKLVKNRTRIV NESDVFSWVIRREFQELRHPVDEEKARCLEGIGG HTRGLVASLDMQLEQAQGTRERLAQAECVLEQF

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GNEDHHEFIWKFHSMASR
3610	A	2	987	DPRVRPPLLQPPPPLLPRLVILKMAPLDLDKYVEIARLCKYLPENDLKRLCDYVCDLLLEESNVQPVSTPVTVCGDIHGQFYDLCELFRTGGQVPDTNYIFMGDFVDRGYYSLETFTYLLALKAKWPDRITLLRGNHESRQITQVYGFYDECQTKYGNANAWRYCTKVFDMLTVAALIDEQILCVHGGLSPDIKTLDQIRTIERNQEIPHKGAFCDLVWSDPEDVDTWAISPRGAGWLFGAKVTNEFVHINNLKLICRAHQLVHEGYKFMFDEKLVTVWSAPNYCYRCGNIASIMVFKDVNTREPKLFRAVPDSERVIPPRTTTPYFL
3611	A	2459	869	AEKMTAELREAMALAPWGPVKVKKEEEEEENFPGQASSQQVHSENIKVWAPVQGLQTGLDGSEEEEKGQNISWDMAVVLKATQEAPAASTLGSYSLPGTLAKSEILETHGTMNFLGAETKNLQLLVPKTEICEEAEKPLIISERIQKADPQGPELGEACEKGNMLKRQRIKREKKDFRQVIVNDCHLPESFKEEENQKCKKSGGKYSLNSGAVKNPKTQLGQKPFTCSVCGKGFSQSANLVVHQRIHTGEKPFECHECGKAFIQSANLVVHQRIHTGQKPYVCSKCGKAFTQSSNLTVHQKIHSLEKTFKCNECEKAFSYSSQLARHQKVHITEKCYECNECGKTFTRSSNLIVHQRIHTGEKPFACNDCGKAFTQSANLIVHQRSHTGEKPYECKECGKAFSCFSHLIVHQRIHTAEKPYDCSECGKAFSQLSCLIVHQRIHSGDLPYVCNECGKAFTCSSYLLIHQRIHNGEKPYTCNECGKAFRQRSSLTVHQRTHTGEKPYECEKCGAAFISNSHLMRHHRTHLVE
3612	A	318	2245	SPMAEAALVNTPQIPMVTEEFVKPSQGHVTFEDIAVYFSQEEWGLLDEAQRCLYHDVMLENFSLMASVGCLHGIEAEEAPSEQTLSAQGVSQARTPKLGPSIPNAHSCEMCILVMKDILYLSEHQGTLPWQKPYTSVASGKWFSFGSNLQQHQNQDSGEKHIRKEESSALLLNSCKIPLSDNLFPCKDVEKDFPTILGLLQHQTTHSRQEYAHRSRETFQQRRYKCEQVFNEKVHVTEHQRVHTGEKAYKRREYGKSLNSKYLFVEHQRTHNAEKPYVCNICGKSFLHKQTLVGHQQRIHTRERSYVCIECGKSLSSKYSLVEHQRTHNGEKPYVCNVCGKSFRHKQTFVGHQQRIHTGERPYVCMECGKSFIHSYDRIRHQRVHTGEGAYQCSECGKSFIYKQSLLDHHRIHTGERPYECKECGKAFIHKKRLLEHQRIHTGEKPYVCIICGKSFIRSSDYMRHQRIHTGERAYECSDCGKAFISKQTLLKHHKIHTRERPYECSECGKGFYLEVKLLQHQRIHTREQLCECNECGKVFSHQKRLLEHQKVHTGEKPCECSECGKCFRHRTSLIQHQKVHSGERPYNCTACEKAFIYKNKLVEHQRIHTGEKPYECGKCGKAFNKRYSLVRHQKVHITEEP
3613	A	817	3345	NQSHPDSETVTVEGGRRKMKSNQERSNECLPPKKREIPATSRSSEEKAPTLPSDNHRVEGTAWLPGNPGGRGHGGGRHGPAGTSVELGLQQGIGLHKALSTGLDYSPPSAPRSVPVATTLPAAYATPQPGTPVSPVQYAHLPHTFQFIGSSQYSGTYASFIPSQLIPPTANPVTSAVASAAGATTPSQRSQLEAYSTLLANMGSLSQTPGHKAEQQQQQQQQQQQQQQQQQQQQQQQQQQHQQQQQQQQQQQQQQHLSRAPGLITPGSPP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PAQQNQYVHISSSPQNTGRTASPPAIPVHLHPHQ TMIPHTLTLGPPSQVVMQYADSGSHFVPREATK KAESSRLQQAIQAKEVLNGEMEKSRRYGAPSSA DLGLGKAGGKSVPHPYESRHVVVHPSPSDYSSR DPSGVRASVMVLPNSNTPAADLEVQQATHREAS PSTLNDKSGLHLGKPGHRSYALSPHTVIQTTHSA SEPLPVGLPATAFYAGTQPPVIGYLSGQQQAITY AGSLPQHLVIPGTQPLLIPVGSTDMEASGAAPAIV TSSPQFAAVPHTFVTTALPKSENFNPEALVTQAA YPAMVQAQIHLPVVQSVASPAAAPPTLPPYFMK GSIIQLANGELKKVEDLKTEDFIQSAEISNDLKIDS STVERIEDSHSPGVAVIQFAVGEHRAQVSVEVLV EYPFFVFGQGWSSCCPERTSQLFDLPCSKLSVGD VCISLTLKNLKNGSVKKGQPVDPASVLLKHSKA DGLAGSRHRYAEQENGINQGSAQMLSENGELKF PEKMGLSAAPFLTKIEPSKPAATRKRRWSAPESR KLEKSEDEPPLTLPKPSLIPQEVKICIEGRSNVGK
3614	A	3	114	FFESRLRCKCCEPRGSWARFGCWRLQPEFKPKQ LEG
3615	A	3	1603	DAWALTNQFSDSKQHIEVLKESLTAKEQRAAILQ TEVDALRLRLEEKETMLNKKTKQIQDMAEEKGT QAGEIHDLKDMLDVKERKVNVLQKKIENLQEQL RDKEKQMSSLKERVKSLQADTTNTDTALTTLEE ALAEKERTIERLKEQRDRDEREKQEEIDNYKKDL KDLKEKVSLLQGDLSEKEASLLDLKEHASSLASS GLKKDSRLKTLEIALEQKKEECLKMESQLKKAH EAALEARASPEMSDRIQHLEREITRYKDESSKAQ AEVDRLLEILKEVENEKNDKDKKIAELESLTSRQ VKDQNKKVANLKHKEQVEKKKSAQMLEEARRR EDNLNDSSQQLQDSLRKKDDRIEELEEALRESVQ ITAEREMVLAQEESARTNAEKQVEELLMAMEKV KQELESMKAKLSSTQQSLAEKETHLTNLRAERR KHLEEVLEMKQEALLAAISEKDANIALLELSSSK KKTQEEVAALKREKDRLVQQLKQQTQNRMKLM ADNYEDDHFKSSHSNQTNHKPSPDQDEEEGIWA
3616	A	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKPDL PTWKRNFRSALNRKEGLRLAEDRSKDPHDPHKI YEFVNSGVGDFSQPDTSPDTNGGGSTSDTQEDIL DELLGNMVLAPLPDPGPPSLAVAPEPCPQPLRSPS LDNPTPFPNLGPSENPLKRLLVPGEEWEFEVTAF YRGRQVFQQTISCPEGLRLVGSEVGDRTLPGWP VTLPDPGMSLTDRGVMSYVRHVLSCLGGGLAL WRAGQWLWAQRLGHCHTYWAVSEELLPNSGH GPDGEVPKDKEGGVFDLGPFIVGSLGPPDLITFTE GSGRSPRYALWFCVGESWPQDQPWTKRLVMVK VVPTCLRALVEMARVGGASSLENTVDLHISNSHP LSLTSDQYKAYLQDLVEGMDFQGPGES
3617	A	852	304	RGGLLSKMARVLKAAAANAVGLFSRLQAPIPTV RASSTSQPLDQVTGSVWNLGRLNHVAIAVPDLE KAAAFYKNILGAQVSEAVPLPEHGVSVVFVNLG NTKMELLHPLGRDSPIAGFLQKNKAGGMHHICIE VDNINAAVMDLKKKKIRSLSEEVKIGAHGKPVIF LHPKDCGGVLVELEQA
3618	A	3	5992	DNIDETYGVNVQFESDEEEGDEDVYGEVREEAS DDDMEGDEAVVRCTLSANMYVDEILVWCASEL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NIPEFFPLESPHKKVGYGLSSRTWLQGGGKVIEA GRDLLVASGELMSSKKKDLHPRDIDAFWLQRQL SRFYDDAIVSQKKADEVLEILKTASDDRECENQL VLLLGFNTFDFIKVLRQHRMMILYCTLLASAQSE AEKERIMGKMEADPELSKFLYQLHETEKEDLIRE ERSRRERVRQSRMDTDLETMDLDQGGEALAPRQ VLDLEDLVFTQGSHFMANKRCQLPDGSFRRQRK GYEEVHVPALKPKPFGSEEQLLPVEKLPKYAQA GFEGFKTLNRIQSKLYRAALETDENLLLCAPTGA GKTNVALMCMLREIGKHINMDGTINVDDFKIIYI APMRSLVQEMVGSFGKRLATYGITVAELTGDHQ LCKEEISATQIIVCTPEKWDIITRKGGERTYTQLV RLIILDEIHLLHDDRGPVLEALVARAIRNIEMTQE DVRLIGLSATLPNYEDVATFLRVDPAKGLFYFDN SFRPVPLEQTYVGITEKKAIKRFQIMNEIVYEKIM EHAGKNQVLVFVHSRKETGKTARAIRDMCLEKD TLGLFLREGSASTEVLRTEAEQCKNLELKDLLPY GFAIHHAGMTRVDRTLVEDLFGDKHIQVLVSTA TLAWGVNLPAHTVIIKGTQVYSPEKGRWTELGA LDILQMLGRAGRPQYDTKGEGILITSHGELQYYL SLLNQQLPIESQMVSKLPDMLNAEIVLGNVQNA KDAVNWLGYAYLYIRMLRSPTLYGISHDDLKGD PLLDQRRLDLVHTAALMLDKNNLVKYDKKTGN FQVTELGRIASHYYITNDTVQTYNQLLKPTLSEIE LFRVFSLSSEFKNITVREEEKLELQKLLERVPIPVK ESIEEPSAKINVLLQAFISQLKLEGFALMADMVY VTQSAGRLMRAIFEIVLNRGWAQLTDKTLNLCK MIDKRMWQSMCPLRQFRKLPEEVVKKIEKKNFP FERLYDLNHNEIGELIRMPKMGKTIHKYVHLFPK LELSVHLQPITRSTLKVELTITPDFQWDEKVHGSS EAFWILVEDVDSEVILHHEYFLLKAKYAQDEHLI TFFVPVFEPLPPQYFIRVVSDRWLSCETQLPVSFR HLILPEKYPPPTELLDLQPLPVSALRNSAFESLYQ DKFPFFNPIQTQVFNTVYNSDDNVFVGAPTGSGK TICAEFAILRMLLQNSEGRCVYITPMRLWQEQVY MDWYEKFQDRLNKKVVLLTGETSTDLKLLGKG NIIISTPEKWDILSRRWKQRKNVQNINLFVVDEV HLIGGENGPVLEVICSRMRYISSQIERPIRIVALSSS LSNAKDVAHWLGCSATSTFNFHPNVRPVPLELHI QGFNISHTQTRLLSMAKPVFHAITKHSPKKPVIVF VPSRKQTRLTAIDILTTCAADIQRQRFLHCTEKDL IPYLEKLSDSTLKETLLNGVGYLHEGLSPMERRL VEQLFSSGAIQVVVASRSLCWGMNVAAHLVIIM DTLYYNGKIHAYVDYPIYDVLQMVGHANRPLQ DDEGRCVIMCQGSKKDFFKKFLYEPLPVESHLD HCMHDHFNAEIVTKTIENKQDAVDYLTWTFLYR RMTQNPNYYNLQGISHRHLSDHLSELVEQTLSDL EQSKCISIEDEMDVAPLNLGMIAAYYYINYTTIEL FSMSLNAKTKVRGLIEIISNAAEYENIPIRHHEDN LLRQLAQKVPHKLNNPKFNDPHVKTNLLLQAHL SRMQLSAELQSDTEEILSKAIRLIQACVDVLSSNG WLSPALAAMELAQMVTQAMWSEDSYLRRLPPF PSGLFKRCTDKGVESVFDIMEMEDEERNALLQLT DSQIADVARFCNRYPNIELSYEVVDKDSIRSGGP VVVLVQLEREEEVTGPVIAPLFPQKREEGWWVV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				IGDAKSNSLISIKRLTLQQKAKVKLDFVAPATGG RHNTLYFMSDAYMGCDQEYKFSVDVKEAETDS DSD
3619	A	3	5992	DNIDETYGVNVQFESDEEEGDEDVYGEVREEAS DDDMEGDEAVVRCTLSANMYVDEILVWCASEL NIPEFFPLESPHKKVGYGLSSRTWLQGGGKVIEA GRDLLVASGELMSSKKKDLHPRDIDAFWLQRQL SRFYDDAIVSQKKADEVLEILKTASDDRECENQL VLLLGFNTFDFIKVLRQHRMMILYCTLLASAQSE AEKERIMGKMEADPELSKFLYQLHETEKEDLIRE ERSRRERVRQSRMDTDLETMDLDQGGEALAPRQ VLDLEDLVFTQGSHFMANKRCQLPDGSFRRQRK GYEEVHVPALKPKPFGSEEQLLPVEKLPKYAQA GFEGFKTLNRIQSKLYRAALETDENLLLCAPTGA GKTNVALMCMLREIGKHINMDGTINVDDFKIIYI APMRSLVQEMVGSFGKRLATYGITVAELTGDHQ LCKEEISATQIIVCTPEKWDIITRKGGERTYTQLV RLIILDEIHLLHDDRGPVLEALVARAIRNIEMTQE DVRLIGLSATLPNYEDVATFLRVDPAKGLFYFDN SFRPVPLEQTYVGITEKKAIKRFQIMNEIVYEKIM EHAGKNQVLVFVHSRKETGKTARAIRDMCLEKD TLGLFLREGSASTEVLRTEAEQCKNLELKDLLPY GFAIHHAGMTRVDRTLVEDLFGDKHIQVLVSTA TLAWGVNLPAHTVIIKGTQVYSPEKGRWTELGA LDILQMLGRAGRPQYDTKGEGILITSHGELQYYL SLLNQQLPIESQMVSKLPDMLNAEIVLGNVQNA KDAVNWLGYAYLYIRMLRSPTLYGISHDDLKGD PLLDQRRLDLVHTAALMLDKNNLVKYDKKTGN FQVTELGRIASHYYITNDTVQTYNQLLKPTLSEIE LFRVFSLSSEFKNITVREEEKLELQKLLERVPIPVK ESIEEPSAKINVLLQAFISQLKLEGFALMADMVY VTQSAGRLMRAIFEIVLNRGWAQLTDKTLNLCK MIDKRMWQSMCPLRQFRKLPEEVVKKIEKKNFP FERLYDLNHNEIGELIRMPKMGKTIHKYVHLFPK LELSVHLQPITRSTLKVELTITPDFQWDEKVHGSS EAFWILVEDVDSEVILHHEYFLLKAKYAQDEHLI TFFVPVFEPLPPQYFIRVVSDRWLSCETQLPVSFR HLILPEKYPPPTELLDLQPLPVSALRNSAFESLYQ DKFPFFNPIQTQVFNTVYNSDDNVFVGAPTGSGK TICAEFAILRMLLQNSEGRCVYITPMRLWQEQVY MDWYEKFQDRLNKKVVLLTGETSTDLKLLGKG NIIISTPEKWDILSRRWKQRKNVQNINLFVVDEV HLIGGENGPVLEVICSRMRYISSQIERPIRIVALSSS LSNAKDVAHWLGCSATSTFNFHPNVRPVPLELHI QGFNISHTQTRLLSMAKPVFHAITKHSPKKPVIVF VPSRKQTRLTAIDILTTCAADIQRQRFLHCTEKDL IPYLEKLSDSTLKETLLNGVGYLHEGLSPMERRL VEQLFSSGAIQVVVASRSLCWGMNVAAHLVIIM DTLYYNGKIHAYVDYPIYDVLQMVGHANRPLQ DDEGRCVIMCQGSKKDFFKKFLYEPLPVESHLD HCMHDHFNAEIVTKTIENKQDAVDYLTWTFLYR RMTQNPNYYNLQGISHRHLSDHLSELVEQTLSDL EQSKCISIEDEMDVAPLNLGMIAAYYYINYTTIEL FSMSLNAKTKVRGLIEIISNAAEYENIPIRHHEDN LLRQLAQKVPHKLNNPKFNDPHVKTNLLLQAHL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SRMQLSAELQSDTEEILSKAIRLIQACVDVLSSNGWLSPALAAMELAQMVTQAMWSEDSYLRRLPPFPSGLFKRCTDKGVESVFDIMEMEDEERNALLQLTDSQIADVARFCNRYPNIELSYEVVDKDSIRSGGPVVVLVQLEREEEVTGPVIAPLFPQKREEGWWVVIGDAKSNSLISIKRLTLQQKAKVKLDFVAPATGGRHNTLYFMSDAYMGCDQEYKFSVDVKEAETDSDSD
3620	A	1205	323	VIKMALAARLLPQFLHSRSLPCGAVRLRTPAVAEVRLPSATLCYFCRCRLGLGAALFPRSARALAASALPAQGSRWPVLSSPGLPAAFASFPACPQRSYSTEEKPQQHQKTKMIVLGFSNPINWVRTRIKAFLIWAYFDKEFSITEFSEGAKQAFAHVSKLLSQCKFDLLEELVAKEVLHALKEKVTSLPDNHKNALAANIDEIVFTSTGDISIYYDEKGRKFVNILMCFWYLTSANIPSETLRGASVFQVKLGNQNVETKQLLSASYEFQREFTQGVKPDWTIARIEHSKLLE
3621	A	2	2995	SSSRSRHSSISPVRLPLNSSLGAELSRKKKERAAAAAAKMDGKESSYERSGSYSGRSPSPYGRRRSSSPFLSKRSLSRSPLPSRKSMKSRSRSPAYSRHSSSHSKKKRSSSRSRHSSISPVRLPLNSSLGAELSRKKKERAAAAAAAKMDGKESSYERSGSYSGRSPSPYGRRRSSSPFLSKRSLSRSPLPSRKSMKSRSRSPAYSRHSSSHSKKKRSSSRSRHSSISPVRLPLNSSLGAELSRKKKERAAAAAAAKMDGKESKGSPVFLPRKENSSVEAKDSGLESKKLPRSVKLEKSAPDTELVNVTHLNTEVKNSSDTGKVKLDENSEKHLVKDLKAQGTRDSKPIALKEEIVTPKETETSEKETPPPLPTIASPPPPLPTTTPPPQTPPLPPLPPIPALPQQPPLPPSQPAFSQVPASSTSTLPPSTHSKTSAVSSQANSQPPVQVSVKTQVSVTAAIPHLKTSTLPPLPLPPLLPGDDDMDSPKETLPSKPVKKEKEQRTRHLLTDLPLPPELPGGDLSPPDSPEPKAITPPQQPYKKRPKICCPRYGERRQTESDWGKRCVDKFDIIGIIGEGTYGQVYKAKDKDTGELVALKKVRLDNEKEGFPITAIREIKILRQLIHRSVVNMKEIVTDKQDALDFKKDKGAFYLVFEYMDHDLMGLLESGLVHFSEDHIKSFMKQLMEGLEYCHKKNFLHRDIKCSNILLNNSGQIKLADFGLARLYNSEESRPYTNKVITLWYRPPKLLLGEERYTPAIDVWSCGCILGELFTKKPIFQANLELAQLELISRLCGSPCPAVWPDVIKLPYFNTMKPKKQYRRRLREEFSFIPSAALDLLDHMLTLDPSKRCTAEQTLQSDFLKDVELSKMAPPDLPHWQDCHELWSKKRRRQRQSGVVVEEPPPSKTSRKETTSGTSTEPVKNSSPAPPQPAPGKVESGAGDAIGLADITQQLNQSELAVLLNLLQSQTDLSIPQMAQLLNIHSNPEMQQQLEALNQSISALTEATSQQQDSETMAPEESLKEAPSAPVILPSAEQTTLEASSTPADMQNILAVLLSQLMKTQEPAGSLEENNSDKNSGPQGPRRTPTMPQEEAAGRSNGGNAL
3622	A	16	390	TPERGSAYPETAAVRRPAGECPITMSDLEAKLSTEHLGDKIKDEDIKLRVIGQDSSEIHFKVKMTTPLKKLKKSYCQRQGVPVNSLRFLFEGQRIADNHTPEELGMEEEDVIEVYQEQIGGHSTV
3623	A	2	1544	PPPAPGPDGLNEGCLHRLSMPHQRPRTCAMNPE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LTMESLGTLHGARGGGSGGGGGGGGGGGGP GHEQELLASPSPHHARRGPRGSLRGPPPPPTAHQ ELGTAAAAAAAASRSAMVTSMASILDGGDYRPE LSIPLHHAMSMSCDSSPPGMGMSNTYTTLTPLQP LPPISTVSDKFHHPHPHHHPHHHHHHHHQRLSGN VSGSFTLMRDERGLPAMNNLYSPYKEMPGMSQS LSPLAATPLGNGLGGLHNAQQSLPNYGPPGHDK MLSPNFDAHHTAMLTRGEQHLSRGLGTPPAAM MSHLNGLHHPGHTQSHGPVLAPSRERPPSSSSGS QVATSGQLEEINTKEVAQRITAELKRYSIPQAIFA QRVLCRSQGTLSDLLRNPKPWSKLKSGRETFRR MWKWLQEPEFQRMSALRLAACKRKEQEPNKDR NNSQKKSRLVFTDLQRRTLFAIFKENKRPSKEMQ ITISQQLGLELTTVSNFFMNARRRSLEKWQDDLS TGGSSSTSSTCTKA
3624	A	27	2152	SARKAEAATSGTAARDGSVGRNLVPPPSASAPK AEVESNEKDNRPEEEEQVIHEDDERPSEKNEFSR RKRSKSEDMDNVQSKRRRYMEEEYEAEFQVKIT AKGDINQKLQKVIQWLLEEKLCALQCAVFDKTL AELKTRVEKIECNKRHKTVLTELQAKIARLTKRF EAAKEDLKKRHEHPPNPPVSPGKTVNDVNSNNN MSYRNAGTVRQMLESKRNVSESAPPSFQTPVNT VSSTNLVTPPAVVSSQPKLQTPVTSGSLTATSVLP APNTATVVATTQVPSGNPQPTISLQPLPVILHVPV AVSSQPQLLQSHPGTLVTNQPSGNVEFISVQSPPT VSGLTKNPVSLPSLPNPTKPNNVPSVPSPSIQRNP TASAAPLGTTLAVQAVPTAHSIVQATRTSLPTVG PSGLYSPSTNRGPIQMKIPISAFSTSSAAEQNSNTT PRIENQTNKTIDASVSKKAADSTSQCGKATGSDS SGVIDLTMDDEESGASQDPKKLNHTPVSTMSSSQ PVSRPLQPIQPAPPLQPSGVPTSGPSQTTIHLLPTA PTTVNVTHRPVTQVTTRLPVPRAPANHQVVYTT LPAPPAQAPLRGTVMQAPAVRQVNPQNSVTVRV PQTTTYVVNNGLTLGSTGPQLTVHHRPPQVHTEP PRPVHPAPLPEAPQPQRLPPEAGSTSRPSEATLEV SHAFRVKMAIVLVMECPGGGSKLCHC
3625	A	210	1115	ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQIT LQGSRRRQGRTAFPASGKKRETDYSDGDPLDVH KRLPSSTGEDRAVMLGFAMMGFSVLMFFLLGTT ILKPFMLSIQREESTCTAIHTDIMDDWLDCAFTCG VHCHGQGKYPCLQVFVNLSHPGQKALLHYNEE AVQINPKCFYTPKCHQDRNDLLNSALDIKEFFDH KNGTPFSCFYSPASQSEDVILIKKYDQMAIFHCLF WPSLTLLGGALIVGMVRLTQHLSLLCEKYSTVV RDEVGGKVPYIEQHQFKLCIMRRSKGRAEKS
3626	A	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVLEG FLSAEECVAMQQRIGEIVAEMDVPLHCRTEFSTQ EEEQLRAQGSTDYFLSSGDKIRFFFEKGVFDEKG NFLVPPEKSINKIGHALHAHDPVFKSITHSFKVQT LARSLGLQMPVVVQSMYIFKQPHFGGEVSPHQD ASFLYTEPLGRVLGVWIAVEDATLENGCLWFIPG SHTSGVSRRMVRAPVGSAPGTSFLGSEPARDNSL FVPTPVQRGALVLIHGEVVHKSKQNLSDRSRQA YTFHLMEASGTTWSPENWLQPTAELPFPQLYT
3627	A	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIPRQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				NPGIFYWIFLPSRSHSASHGSRQRQVSCQGTQDEI LKMRNTFAELKNSLEALSSRMDQAEERIGTQAG VQWRDHGSLQPQPPEFKQCFHLSLPSSWDYRAC LS
3628	A	2	810	GCKHLLQNSWYDPRVREADRVGQRARRPRAAM DWLMGKSKAKPNGKKPAAEERKAYLEPEHTKA RITDFQFKELVVLPREIDLNEWLASNTTTFFHHIN LQYSTISEFCTGETCQTMAVCNTQYYWYDERGK KVKCTAPQYVDFVMSSVQKLVTDEDVFPTKYG REFPSSFESLVRKICRHLFHVLAHIYWAHFKETLA LELHGHLNTLYVHFILFAREFNLLDPKETAIMDD LTEVLCSGGRRGSTVGAVGMGPAAGAPGAQNH VKER
3629	A	699	1604	CSHGSSAVSAWSPLFQASEVERQLSMQVHALRE DFREKNSSTNQHIIRLESLQAEIKMLSDRKRELEH RLSATLEENDLLQGTVEELQDRVLILERQGHDKD LQLHQSQLELQEVRLSCRQLQVKVEELTEERSLQ SSAATSTSLLSEIEQSMEAEELEQEREQLTLLSVE MTALKEERDRLRVTSEDKEPKEQLQKAIRDRDE AIAKKNAVELELAKCRMDMMSLNSQLLDAIQQ KLNLSQQLEAWQDDMHRVIDRQLMDTHLKERS QPAAALCRGHSAGRGDEPSIAEGKRLFSFFRKI
3630	A	423	1	PAKVLTLDIYLSKTEGAQVDEPVVITPRAEDCGD WDDMEKRSSGRRSGRRRGSQKSTDSPGADAELP ESAARDDAVFDDEVAPNAASDNASAEKKVKSPR AALDGGVASAASPESKPSPGTKGQLRGESDRSK QPPPASSP
3631	A	2082	674	WSGFWQLPGVRGVGSAPGGDGAEFTSRRGSSRR PGAACPGCRGAGSERAPGGMGRRRAPELYRAPF PLYALQVDPSTGLLIAAGGGGAAKTGIKNGVHF LQLELINGRLSASLLHSHDTETRATMNLALAGDI LAAGQDAHCQLLRFQAHQQQGNKAEKAGSKEQ GPRQRKGAAPAEKKCGAETQHEGLELRVENLQA VQTDFSSDPLQKVVCFNHDNTLLATGGTDGYVR VWKVPSLEKVLEFKAHEGEIEDLALGPDGKLVT VGRDLKASVWQKDQLVTQLHWQENGPTFSSTP YRYQACRFGQVPDQPAGLRLFTVQIPHKRLRQPP PCYLTAWDGSNFLPLRTKSCGHEVVSCLDVSES GTFLGLGTVTGSVAIYIAFSLQCLYYVREAHGIV VTDVAFLPEKGRGPELLGSHETALFSVAVDSRCQ LHLLPSRRSVPVWLLLLLCVGLIIVTILLLQSAFPG FL
3632	A	942	40	PWCQRVEVRSCGSSKRSCSRWSGSSWDGSRSLG RGLNHTSLNRSPPFTPDTMTHCCSPCCQPTCCRT TCCRTTCWKPTTVTTCSSTPCCQPSCCVPSCCQP CCHPTCCQNTCCRTTCCQPTCVASCCQPSCCSTP CCQPTCCGSSCCGQTSCGSSCCQPICGSSCCQPCC HPTCYQTICFRTTCCQPTCCQPTCCRNTSCQPTCC GSSCCQPCCHPTCCQTICRSTCCQPSCVTRCCSTP CCQPTCGGSSCCSQTCNESSYCLPCCRPTCCQTT CYRTTCCRPSCCCSPCCVSSCCQPSCC
3633	A	605	3004	GPEGYRGRRARHPSLGSTTGHCGGGRGAEGTGT DPAAPAARLNVDGLLVYFPYDYIYPEQFSYMRE LKRTLDAKGHGVLEMPSGTGKTVSLLALIMAYQ RAYPLEVTKLIYCSRTVPEIEKVIEELRKLLNFYE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KQEGEKLPFLGLALSSRKNLCIHPEVTPLRFGKD VDGKCHSLTASYVRAQYQHDTSLPHCRFYEEFD AHGREVPLPAGIYNLDDLKALGRRQGWCPYFLA RYSILHANVVVYSYHYLLDPKIADLVSKELARK AVVVFDEAHNIDNVCIDSMSVNLTRRTLDRCQG NLETLQKTVLRIKETDEQRLRDEYRRLVEGLREA SAARETDAHLANPVLPDEVLQEAVPGSIRTAEHF LGFLRRLLEYVKWRLRVQHVVQESPPAFLSGLA QRVCIQRKPLRFCAERLRSLLHTLEITDLADFSPL TLLANFATLVSTYAKGFTIIIEPFDDRTPTIANPIL HFSCMDASLAIKPVFERFQSVIITSGTLSPLDIYPK ILDFHPVTMATFTMTLARVCLCPMIIGRGNDQVA ISSKFETREDIAVIRNYGNLLLEMSAVVPDGIVAF FTSYQYMESTVASWYEQGILENIQRNKLLFIETQ DGAETSVALEKYQEACENGRGAILLSVARGKVS EGIDFVHHYGRAVIMFGVPYVYTQSRILKARLEY LRDQFQIRENDFLTFDAMRHAAQCVGRAIRGKT DYGLMVFADKRFARGDKRGKLPRWIQEHLTDA NLNLTVDEGVQVAKYFLRQMAQPFHREDQLGL SLLSLEQLESEETLKRIEQIAQQL
3634	A	159	384	LKMSSKTASTNNIAQARRTVQQLRLEASIERIKV SKASADLMSYCEEHARSDPLLIGIPTSENPFKDKK TCIIL
3635	A	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMTGT LETQFTCPFCNHEKSCDVKMDRARNTGVISCTV CLEEFQTPITCILGNLGFFQRVGRGLESGPCSSGP LCALVQGQSRPEEQVPPSDFCGVRRCRAGFQCQ
3636	A	48	282	DHLKSCYQDSHEDPTKMKRFLFLLLTISLLVMVQ IQTGLSGQNDTSQTSSPSASSSMSGGIFLFFVANAI IHLFCFS
3637	A	1	1248	ARAGSVVGSAAARGPPAGCRCERAARLPSSPAR RRRCDWVEDGAGRMEILMTVSKFASICTMGAN ASALEKEIGPEQFPVNEHYFGLVNFGNTCYCNSV LQALYFCRPFREKGLAYKSQPRKKESLLTCLADL FHSIATQKKKVGVIPPKKFITRLRKENELFDNYM QQDAHEFLNYLLNTIADILQEERKQEKQNGRLPN GNIDNENNNSTPDPTWVHEIFQGTLTNETRCLTC ETISSKDEDFLDLSVDVEQNTSITHCLRGFSNTET LCSEYKYYCEECRSKQEAHKRMKVKKLPMILAL HLKRFKYMDQLHRYTKLSYRVVFPLELRLFNTS GDATNPDRMYDLVAVVVHCGSGPNRGHYIAIV KSHDFWLLFDDDIVEKIDAQAIEEFYGLTSDISKN SESGYILFYQSRD
3638	A	11	630	PAGIPVSTISSDRRASTDLTRKMKPDETPMFDPNL LKEVDWSQNTATFSPAISPTHPGEGLVLRPLCTA DLNRGFFKVLGQLTETGVVSPEQFMKSFEHMKK SGDYYVTVVEDVTLGQIVATATLIIEHKFIHSCAK RGRVEDVVVSDECRGKQLGNLLLSTLTLLSKKL NCYKITLECLPQNVGFYKKFGYTVSEENYMCRR FLK
3639	A	2	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHSSPL LPHAMKSPFYRCQNTTSVEKGNSAVMGGVLFST GLLGNLLALGLLARSGLGWCSRRPLRPLPSVFY MLVCGLTVTDLLGKCLLSPVVLAAYAQNRSLRV LAPALDNSLCQAFAFFMSFFGLSSTLQLLAMALE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				CWLSLGHPFFYRRHITLRLGALVAPVVSAFSLAF CALPFMGFGKFVQYCPGTWCFIQMVHEEGSLSV LGYSVLYSSLMALLVLATVLCNLGAMRNLYAM HRRLQRHPRSCTRDCAEPRADGREASPQPLEELD HLLLLALMTVLFTMCSLPVIYRAYYGAFKDVKE KNRTSEEAEDLRALRFLSVISIVDPWIFIIFRSPVFR IFFHKIFIRPLRYRSRCSNSTNMESSL
3640	A	930	182	PLPPPTLAMFLTRSEYDRGVNTFSPEGRLFQVEY AIEAIKLGSTAIGIQTSEGVCLAVEKRITSPLMEPS SIEKIVEIDAHIGCAMSGLIADAKTLIDKARVETQ NHWFTYNETMTVESVTQAVSNLALQFGEEDADP GAMSRPFGVALLFGGVDEKGPQLFHMDPSGTFV QCDARAIGSASEGAQSSLQEVYHKSMTLKEAIKS SLIILKQVMEEKLNATNIELATVQPGQNFHMFTK EELEEVIKDI
3641	A	2	1254	PTGQGGRRAEARSCLLSKAMLGRSGYRALPLGD FDRFQQSSFGFLGSQKGCLSPERGGVGTGADVPQ SWPSCLCHGLISFLGFLLLLVTFPISGWFALKIVPT YERMIVFRLGRIRTPQGPGMVLLLPFIDSFQRVDL RTRAFNVPPCKLASKDGAVLSVGADVQFRIWDP VLSVMTVKDLNTATRMTAQNAMTKALLKRPLR EIQMEKLKISDQLLLEINDVTRAWGLEVDRVELA VEAVLQPPQDSPAGPNLDSTLQQLALHFLGGSM NSMAGGAPSPGPADTVEMVSEVEPPAPQVGARS SPKQPLAEGLLTALQPFLSEALVSQVGACYQFNV VLPSGTQSAYFLDLTTGRGRVGHGVPDGIPDVV VEMAEADLRALLCRELRPLGAYMSGRLKVKGD LAMAMKLEAVLRALK
3642	A	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPRKH DSGAADLERVTDYAEEKEIQSSNLETAMSVIGDR RSREQKAKQER
3643	A	94	541	RKERRRRRRRMEAVVFVFSLLDCCALIFLSVYFII TLSDLECDYINARSCCSKLNKWVIPELIGHTIVTV LLLMSLHWFIFLLNLPVATWNIYRYIMVPSGNM GVFDPTEIHNRGQLKSHMKEAMIKLGFHLLCFF MYLYSMILALIND
3644	A	95	2808	TSCRHFPITSEDPLNYLLILTVERIYAYQALPLGFL FCSRDPVPEYLNHCGVKYVLISDRASFCALHIFFS PFRNVFRPAAGGGIAPPPRLWFQPSLSDAEMEIPK LLPARGTLQGGGGGGIPAGGGRVHRGPDSPAGQ VPTRRLLLPRGPQDGGPGRRREEASTASRGPGPS LFAPRPHQPSGGGGGGGDDFFLVLLDPVGGDVE TAGSGQAAGPVLREEAEEGPGLQGGESGANPAG PTALGPRCLSAVPTPAPISAPGPAAAFAGTVTIHN QDLLLRFENGVLTLATPPPHAWEPGAAPAQQPG CLIAPQAGFPHAAHPGDCPELPPDLLLAEPAEPAP APAPEEEAEGPAAALGPRGPLGSGPGVVLYLCPE ALCGQTFAKKHQLKMHLLTHSSSQGQRPFKCPL GGCGWTFTTSYKLKRHLQSHDKLRPFGCPAEGC GKSFTTVYNLKAHMKGHEQENSFKCEVCEESFP TQAKLGAHQRSHFEPERPYQCAFSGCKKTFITVS ALFSHNRAHFREQELFSCSFPGCSKQYDKACRLK IHLRSHTGERPFLCDFDGCGWNFTSMSKLLRHKR KHDDDRRFMCPVEGCGKSFTRAEHLKGHSITHL STKPFVCPVAGCCARFSARSSLYIHSKKHLQDVD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TWKSRCPISSCNKLFTSKHSMKTHMVKRHKVGQ DLLAQLEAANSLTPSSELTSQRQNDLSDAEIVSLF SDVPDSTSAALLDTALVNSGILTIDVASVSSTLAG HLPANNNNSVGQAVDPPSLMATSDPPQSLDTSLF FGTAATGFQQSSLNMDEVSSVSVGPLGSLDSLA MKNSSPEPQALTPSSKLTVDTDTLTPSSTLCENSV SELLTPAKAEWSVHPNSDFFGQEGETQFGFPNAA GNHGSQKERNLITVTGSSFLV
3645	A	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPACRV PYCSVVCFRKHKEQCNPETRPVEKKIRSALPTKT VKPVENKDDDDSIADFLNSDEEEDRVSLQNLKN LGESATLRSLLLNPHLRQLMVNLDQGEDKAKLM RAYMQEPLFVEFADCCLGIVEPSQNEES
3646	A	85	1948	ERGGGKAAAAAAAAAAARALAASGQDPRPHPR APPWDDSGDDDEATTPADKSELHHTLKNLSLKL DDLSTCNDLIAKHGAALQRSLTELDGLKIPSESG EKLKVVNERATLFRITSNAMINACRDFLELAEIHS RKWQRALQYEQEQRVHLEETIEQLAKQHNSLER AFHSAPGRPANPSKSFIEGSLLTPKGEDSEEDEDT EYFDAMEDSTSFITVITEAKEDSRKAEGSTGTSSA DWSSADNVLDGASLVPKGSSKVKRRVRIPNKPN YSLNLWSIMKNCIGRELSRIPMPVNFNEPLSMLQ RLTEDLEYHHLLDKAVHCTSSVEQMCLVAAFSV SSYSTTVHRIAKPFNPMLGETFELDRLDDMGLRS LCEQVSHHPPSAAHYVFSKHGWSLWQEITISSKF RGKYISIMPLGAIHLEFQASGNHYVWRKSTSTVH NIIVGKLWIDQSGDIEIVNHKTNDRCQLKFLPYSY FSKEAARKVTGVVSDSQGKAHYVLSGSWDEQM ECSKVMHSSPSSPSSDGKQKTVYQTLSAKLLWK KYPLPENAENMYYFSELALTLNEHEEGVAPTDS RLRPDQRLMEKGRWDEANTEKQRLEEKQRLSR RRRLEACGPGSSCSSEE
3647	A	46	5007	PTGDACVSTSCELASALSHLDASHLTENLPKAAS ELGQQPMTELDSSSDLISSPGKKGAAHPDPSKTS VDTGQVSRPENPSQPASPRVTKCKARSPVRLPHE GSPSPGEKAAAPPDYSKTRSASETSTPHNTRRVA ALRGAGPGAEGMTPAGAVLPGDPLTSQEQRQGA PGNHSKALEMTGIHAPESSQEPSLLEGADSVSSR APQASLSMLPSTDNTKEACGHVSGHCCPGGSRE SPVTDIDSFIKELDASAARSPSSQTGDSGSQEGSA QGHPPAGAGGGSSCRAEPVPGGQTSSPRRAWAA GAPAYPQWASQPSVLDSINPDKHFTVNKNFLSN YSRNFSSFHEDSTSLSGLGDSTEPSLSSMYGDAE DSSSDPESLTEAPRASARDGWSPPRSRVSLHKED PSESEEEQIEICSTRGCPNPPSSPAHLPTQAAICPAS AKVLSLKYSTPRESVASPREKVACLPGSYTSGPD SSQPSSLLEMSSQEHETHADISTSQNHRPSCAEET TEVTSASSAMENSPLSKVARHFHSPPIILSSPNMV NGLEHDLLDDETLNQYETSINAAASLSSFSVDVP KNGESVLENLHISESQDLDDLLQKPKMIARRPIM AWFKEINKHNQGTHLRSKTEKEQPLMPARSPDS KIQMVSSSQKKGVTVPHSPPQPKTNLENKDLSKK SPAEMLLTNGQKAKCGPKLKRLSLKGKAKVNSE APAANAVKAGGTDHRKPLISPQTSHKTLSKAVS QRLHVADHEDPDRNTTAAPRSPQCVLESKPPLAT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SGPLKPSVSDTSIRTFVSPLTSPKPVPEQGMWSRF HMAVLSEPDRGCPTTPKSPKCRAEGRAPRADSG PVSPAASRNGMSVAGNRQSEPRLASHVAADTAQ PRPTGEKGGNIMASDRLERTNQLKIVEISAEAVSE TVCGNKPAESDRRGGCLAQGNCQEKSEIRLYRQ VAESSTSHPSSLPSHASQAEQEMSRSFSMAKLAS SSSSLQTAIRKAEYSQGKSSLMSDSRGVPRNSIPG GPSGEDHLYFTPRPATRTYSMPAQFSSHFGREGH PPHSLGRSRDSQVPVTSSVVPEAKASRGGLPSLA NGQGIYSVKPLLDTSRNLPATDEGDIISVQETSCL VTDKIKVTRRHYCYEQNWPHESTSFFSVKQRIKS FENLANADRPVAKSGASPFLSVSSKPPIGRRSSGS IVSGSLGHPGDAAARLLRRSLSSCSENQSEAGTL LPQMAKSPSIMTLTISRQNPPETSSKGSDSELKKS LGPLGIPTPTMTLASPVKRNKSSVRHTQPSPVSRS KLQELRALSMPDLDKLCSEDYSAGPSAVLFKTEL EITPRRSPGPPAGGVSCPEKGGNRACPGGSGPKT SAAETPSSASDTGEAAQDLPFRRSWSVNLDQLLV SAGDQQRLQSVLSSVGSKSTILTLIQEAKAQSENE EDVCFIVLNRKEGSGLGFSVAGGTDVEPKSITVH RVFSQGAASQEGTMNRGDFLLSVNGASLAGLAH GNVLKVLHQAQLHKDALVVIKKGMDQPRPSAR QEPPTANGKGLLSRKTIPLEPGIGRSVAVHDALC VEVLKTSAGLGLSLDGGKSSVTGDGPLVIKRVY KGGAAEQAGIIEAGDEILAINGKPLVGLMHFDA WNIMKSVPEGPVQLLIRKHRNSS
3648	A	337	1564	KSRLSVTLMPVQLSEHPEWNESMHSLRISVGGLP VLASMTKAADPRFRPRWKVVLTFFVGAAILWLL CSHRPAPGRPPTHNAHNWRLGQAPANWYNDTY PLSPPQRTPAGIRYRIAVIADLDTESRAQEENTWF TYLKKGYLTFSDSGDKVAVEWDKDHGVLESHL AEKGRGMELSDLIVFNGKLYSVDDRTGVVYQIE GSKAVPWVILSDGDGTVEKGFKAEWLAVKDER LYVGGLGKEWTTTTGDVVNENPEWVKVVGYK GSVDHENWVSNYNALRAAAGIQPPGYLIHESAC WSDTLQRWFFLPRRASQERYSEKDDERKGANLL LSASPDFGDIAVSHVGAVVPTHGFSSFKFIPNTDD QIIVALKSEEDSGRVASYIMAFTLDGRFLLPETKI GSVKYEGIEFI
3649	A	1	775	PTRPGSGSAGGARVGSGEFGVEMAALAPLPPLPA QFKSIQHHLRTAQEHDKRDPVVAYYCRLYAMQ TGMKIDSKTPECRKFLSKLMDQLEALKKQLGDN EAITQEIVGCAHLENYALKMFLYADNEDRAGRF HKNMIKSFYTASLLIDVITVFGELTDENVKHRKY ARWKATYIHNCLKNGETPQAGPVGIEEDNDIEEN EDAGAASLPTQPTQPSSSSTYDPSNMPSGNYTGI QIPPGAHAPANTPAEVPHSTGVAK
3650	A	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLLLL LLGSGQGPQQVGAGQTFEYLKREHSLSKPYQGE APRPCFLRDWELQVHFKIHGQGKKNLHGDGLAI WYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNE EKQQERVFPYISAMVNNGSLSYDHERDGRPTEL GGCTAIVRNLHYDTFLVIRYVKRHLTIMMDIDGK HEWRDCIEVPGVRLPRGYYFGTSSITGDLSDNHD VISLKLFELTVERTPEEEKLHRDVFLPSVDNMKL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PEMTAPLPPLSGLALFLIVFFSLVFSVFAIVIGIILY NKWQEQSRKRFY
3651	A	1	1218	RSWAYVKKCKNNMCPNRGLHDGPEPCWLHHA AGTVSAVQARGLQPSQSRSRPRVPGLATALAYG PAHTPPLSRIGWAMQPPPPGPLGDCLRDWEDLQ QDFQNIQVSAAADAGSPPSRVSLAQGQGSGSPGC KPSLPAEAEGAAQELENQMKERQGLFFDMEAYL PKKNGLYLSLVLGNVNVTLLSKQAKFAYKDEYE KFKLYLTIILILISFTCRFLLNSRVTDAAFNFLLVW YYCTLTIRESILINNGSRIKGWWVFHHYVSTFLSG VMLTWPDGLMYQKFRNQFLSFSMYQSFVQFLQ YYYQSGCLYRLRALGERHTMDLTVEGFQSWMW RVLTFLLPFLFFGHFWQLFNALTLFNLAQDPQCK EWQVLMCGFPFLLLFLGNFFTTLRVVHHKFHSQ RHGSKKD
3652	A	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQPM MQTIGQKYCMDPAVIAGVLSRKSPGDKILVNMG DRTSMVQDPGSQAPTSWISESQVFQTTEVLTTRI TELQRRFPTWTPDQYLRGGLCAYSGGAGYVRSS QDLSCDFCNDVLARAKYLKRHGF
3653	A	2	909	IVRRDWQEVSDIHLAMANCKMTKSIRFPALEHC YTGGEVVLPKDQEEWKRRTGLLLYENYGQSETG LICATYWGMKIKPGFMGKATPPYDVQFHMEASV ENCIIVSMNTADPGSQGITHSLLLQVIDDKGSILPP NTEGNIGIRIKPVRPVSLFMCYEGDPEKTAKVEC GDFYNTGDRGKMDEEGYICFLGRSDDIINASGYR IGPAEVESALVEHPAVAESAVVGSPDPIRGEVVK AFIVLTPQFLSHDKDQLTKELQQHVKSVTAPYKY PRKVEFVSELPKTITGKIERKELRKKETGQM
3654	A	2	909	IVRRDWQEVSDIHLAMANCKMTKSIRFPALEHC YTGGEVVLPKDQEEWKRRTGLLLYENYGQSETG LICATYWGMKIKPGFMGKATPPYDVQFHMEASV ENCIIVSMNTADPGSQGITHSLLLQVIDDKGSILPP NTEGNIGIRIKPVRPVSLFMCYEGDPEKTAKVEC GDFYNTGDRGKMDEEGYICFLGRSDDIINASGYR IGPAEVESALVEHPAVAESAVVGSPDPIRGEVVK AFIVLTPQFLSHDKDQLTKELQQHVKSVTAPYKY PRKVEFVSELPKTITGKIERKELRKKETGQM
3655	A	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDTG MVAHINNSRLKAKGVGQHDNAQNFGNQSFEEL RAACLRKGELFEDPLFPAEPSSLGFKDLGPNSKN VQNISWQRPKDIINNPLFIMDGISPTDICQGILGDC WLLAAIGSLTTCPKLLYRVVPRGQSFKKNYAGIF HFQIWQFGQWVNVVVDDRLPTKNDKLVFVHST ERSEFWSALLEKAYAKLSGSYEALSGGSTMEGL EDFTGGVAQSFQLQRPPQNLLRLLRKAVERSSL MGCSIEVTSDSELESMTDKMLVRGHAYSVTGLQ DVHYRGKMETLIRVRNPWGRIEWNGAWSDSAR EWEEVASDIQMQLLHKTEDGEFWMSYQDFLNN FTLLEICNLTPDTLSGDYKSYWHTTFYEGSWRTG SSAGGCRNHPGTFWTNPQFKISLPEGDDPEDDAE GNVVVCTCLVALMQKNWRHARQQGAQLQTIGF VLYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEIF TNSREVSSQLRLPPGEYIIIPSTFEPHRDADFLLRV FTEKHSESWELDEVNYAEQLQEEKVSEDDMDQ

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DFLHLFKIVAGEGKEIGVYELQRLLNRMAIKFKS FKTKGFGLDACRCMINLMDKDGSGKLGLLEFKI LWKKLKKWMDIFRECDQDHSGTLNSYEMRLVIE KAGIKLNNKVMQVLVARYADDDLIIDFDSFISCF LRLKTMFTFFLTMDPKNTGHICLSLEQVLGEGW EGICRIAPACPSTPPPPSSDVPGPASCPRLFPPWDL LPVSTVAADDHVGIEAL
3656	A	3	174	PLCTHYLLPELPEKSSRTSPRSRPGNMLSGDPHLP QPLCHCLDHCPCCFSGKRLVA
3657	A	1	444	DTRSTYHNAHSLPTYVKSPAPCQMTYIKSPAPCQ TQTCYVQGASPCQSYYVQAPASGSTSQYCVTDP CSAPCSTSYCCLAPRTFGVSPLRRWIQRPQNCNT GSSGCCENSGSSGCCGSGGCGCSCGCGSSGCCCL GIIPMKSRSPALL
3658	A	92	1537	SEAPVQPQPYTMTSFYSTSSCPLGCTMAPGARNV FVSPIDVGCQPVAEANAASMCLLANVAHANRVR VGSTPLGRPSLCLPPTSHTACPLPGTCHIPGNIGIC GAYGKNTLNGHEKETMKFLNDRLANYLEKVRQ LEQENAELETTLLERSKCHESTVCPDYQSYFRTIE ELQQKILCSKAENARLIVQIDNAKLAADDFRIKL ESERSLHQLVEADKCGTQKLLDDATLAKADLEA QQESLKEEQLSLKSNHEQEVKILRSQLGEKFRIEL DIEPTIDLNRVLGEMRAQYEAMVETNHQDVEQ WFQAQSEGISLQAMSCSEELQCCQSEILELRCTV NALEVERQAQHTLKDCLQNSLCEAEDRYGTELA QMQSLISNLEEQLSEIRADLERQNQEYQVLLDVK ARLENEIATYRNLTPLQSLFHACLLYFLSKLWPC HRWVSLWPWSQHGEMILKARVRRLRLVALGSG VPSPCPVFLQD
3659	A	2	402	DLLQCLNQLYSASTEMSCQQSQQQCQPPPKCTP KCPPKCTPKCPPKCPPKCPPQYSAPCPPPVSSCCG SSSGGCCSSEGGGCCLSHHRPRQSLRRRPQSSSC CGSGSGQQSGGSSCCHSSGGSGCCHSSGGCC
3660	A	26	710	CSAVEVKMAARTAFGAVCRRLWQGLGNFSVNT SKGNTAKNGGLLLSTNMKWVQFSNLHVDVPKD LTKPVVTISDEPDILYKRLSVLVKGHDKAVLDSY EYFAVLAAKELGISIKVHEPPRKIERFTLLQSVHI YKKHRVQYEMRTLYRCLELEHLTGSTADVYLEY IQRNLPEGVAMEVTKFCFFIFLDTIRTVTRTHQGA NLGNTIRRKRRKQVIKPQGGHFCLNLK
3661	A	2	370	DVSVAASEPTVYRNPTKMSCQQNQQQCQPPPKC PIPKYPPKCPSKCASSCPPPISSCCGSSSGGCCSSG GCGCCSSEGGGCCLSHHRHHRSHCHRPKSSNCY GSGSGQQSGGSGCCSGGGCC
3662	A	205	1277	RKSLPHPNPQKMLKKPLSAVTWLCIFIVAFVSHP AWLQKLSKHKTPAQPQLKAANCCEEVKELKAQ VANLSSLLSELNKKQERDWVSVVMQVMELESN SKRMESRLTDAESKYSEMNNQIDIMQLQAAQTV TQTSAGKETSPLRERGVPPHLQHCFYIPPDDFLGS PELEVFCDMETSGGGWTIIQRRKSGLVSFYRDW KQYKQGFGSIRGDFWLGNEHIHRLSRQPTRLRVE MEDWEGNLRYAEYSHFVLGNELNSYRLFLGNY TGNVGNDALQYHNNTAFSTKDKDNDNCLDKCA QLRKGGYWYNCCTDSNLNGVYYRLGEHNKHLD GITWYGWHGSTYSLKRVEMKIRPEDFKP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3663	A	64	1456	LSSAKETLAQMYNTVWNMEDLDLEYAKTDINC GTDLMFYIEMDPPALPPKPPKPTTVANNGMNNN MSLQDAEWYWGDISREEVNEKLRDTADGTFLV RDASTKMHGDYTLTLRKGGNNKLIKIFHRDGKY GFSDPLTFSSVVELINHYRNESLAQYNPKLDVKL LYPVSKYQQDQVVKEDNIEAVGKKLHEYNTQFQ EKSREYDRLYEEYTRTSQEIQMKRTAIEAFNETIK IFEEQCQTQERYSKEYIEKFKREGNEKEIQRIMHN YDKLKSRISEIIDSRRRLEEDLKKQAAEYREIDKR MNSIKPDLIQLRKTRDQYLMWLTQKGVRQKKL NEWLGNENTEDQYSLVEDDEDLPHHDEKTWNV GSSNRNKAENLLRGKRDGTFLVRESSKQGCYAC SVVVDGEVKHCVINKTATGYGFAEPYNLYSSLK ELVLHYQHTSLVQHNDSLNVTLAYPVYAQQRR
3664	A	944	406	GATVEDQSCNFGSLRWVVSVPHISARSCPDPLLS RTGRVPGGRGAGLPRHHSPRCCLQVFFNGANVR QVDVPTLTGAFGILAAHVPTLQVLRPGLVVVHA EDGTTSKYFVSSGSIAVNADSSVQLLAEEAVTLD MLDLGAAKANLEKAQAELVGTADEATRAEIQIR IEANEALVKALE
3665	A	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSFHEH RHQSGRCLSTGMAPNLKGRPRKKKPCPQRRDSF SGVKDSNNNSDGKAVAKVKCEARSALTKPKNN HNCKKVSNEEKPKVAIGEECRADEQAFLVALYK YMKERKTPIERIPYLGFKQINLWTMFQAAQKLG GYETITARRQWKHIYDELGGNPGSTSAATCTRR HYERLILPYERFIKGEEDKPLPPIKPRKQENSSQE NENKTKVSGTKRIKHEIPKSKKEKENAPKPQDAA EVSSEQEKEQETLISQKSIPEPLPAADMKKKIEGY QEFSAKPLASRVDPEKDNETDQGSNSEKVAEEA GEKGPTPPLPSAPLAPEKDSALVPGASKQPLTSPS ALVDSKQESKLCCFTESPESEPQEASFPRLPHHTG HRWQTRMRRRMTNCPPWQITLPTAP
3666	A	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTIYPGI KARITQRALDYGVQAGMKMIEQMLKEKKLPDL SGSESLEFLKVDYVNYNFSNIKISAFSFPNTSLAF VPGVGIKALTNHGTANISTDWGFESPLFVLYNSF AEPMEKPILKNLNEMLCPIIASEVKALNANLSTLE VLTKIDNYTLLDYSLISSPEITENYLDLNLKGVFY PLENLTDPPFSPVPFVLPERSNSMLYIGIAEYFFKS ASFAHFTAGVFNVTLSTEEISNHFVQNSQGLGNV LSRIAEIYILSQPFMVRIMATEPPIINLQPGNFTLDI PASIMMLTQPKNSTVETIVSMDFVASTSVGLVIL GQRLVCSLSLNRFRLALPESNRSNIEVLRFENILSS ILHFGVLPLANAKLQQGFPLPNPHKFLFVNSDIEV LEGFLLISTDLKYETSSKQQPSFHVWEGLNLISRQ WRGKSAP
3667	A	1	181	FRGRLGSGRNGGGSMNAPPAFESFLLFEGEKITIN KDTKVPNACLFTINKEDHTLGNIIK
3668	A	212	431	VAGEAVPFFPMMYSEPLKPSYLALVLWYFLLTG YCITKPEVIFKIEQGEEPWILEKGFPSQCHPAKYL WCLHD
3669	A	458	1056	FSGVCFAGIAGSMATLLHDAVMNPAEVVKQRLQ MYNSQHRSAISCIRTVWRTEGLGAFYRSYTTQLT MNIPFQSIHFITYEFLQEQVNPHRTYNPQSHIISGG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LAGALAAAATTPLDVCKTLLNTQENVALSLANIS GRLSGMANAFRTVYQLNGLAGYFKGIQARVIYQ MPSTAISWSVYEFFKYFLTKRQLENRAPY
3670	A	145	298	RNPCPLTFLPSTLMVLLLSLTFFSALTFHSICQLRN TGVEVDIVFQRVSFL
3671	A	3	462	ILKVAKKERTMSSLPVPYKLPVSLSVGSCVIIKGT PIHSFINDPQLQVDFYTDMDEDSDIAFRFRVHFG NHVVMNRREFGIWMLEETTDYVPFEDGKQFELC IYVHYNEYEIKVNGHTHLRALSHRIPPSFVEDGC KCPRRYLPWTSVCVCN
3672	A	1	1028	HYAKLGTRPRLKFMSSPSLSDLGKREPAAAADE RGTQQRRACANATWNSIHNGVIAVFQRKGLPDQ ELFSLNEGVRQLLKTELGSFFTEYLQNQLLTKGM VILRDKIRFYEGQKLLDSLAETWDFFFSDVLPML QAIFYPVQGKEPSVRQLALLHFRNAITLSVKLED ALARAHARVPPAIVQMLLVLQGVHESRGVTEDY LRLETLVQKVVSPYLGTYGLHSSEGPFTHSCILEK RLLRRSRSGDVLAKNPVVRSKSYNTPLLNPVQE HEAEGAAAGGTSIRRHSVSEMTSCPEPQGFSDPP GQGPTGTFRSSPAPHSGPCPSRLYPTTQPPEQGLD PTRS
3673	A	2	712	RPPRVWYPELRELSAAAPRWSHRTAPGIMVFYF TSSSVNSSAYTIYMGKDKYENEDLIKHGWPEDI WFHVDKLSSAHVYLRLHKGENIEDIPKEVLMDC AHLVKANSIQGCKMNNVNVVYTPWSNLKKTAD MDVGQIGFHRQKDVKIVTVEKKVNEILNRLEKT KVERFPDLAAEKECRDREERNEKKAQIQEMKKR EKEEMKKKREMDELRSYSSLMKVENMSSNQDG NDSDEFM
3674	A	2	712	RPPRVWYPELRELSAAAPRWSHRTAPGIMVFYF TSSSVNSSAYTIYMGKDKYENEDLIKHGWPEDI WFHVDKLSSAHVYLRLHKGENIEDIPKEVLMDC AHLVKANSIQGCKMNNVNVVYTPWSNLKKTAD MDVGQIGFHRQKDVKIVTVEKKVNEILNRLEKT KVERFPDLAAEKECRDREERNEKKAQIQEMKKR EKEEMKKKREMDELRSYSSLMKVENMSSNQDG NDSDEFM
3675	A	921	1321	VTLAKMRVHISSCLKVQEQMANCPKFVPVVPTS QPIPSNIPNRSTFACPYCGARNLDQQELVKHCVE SHRSDPNRVVCPICSAMPWGDPSYKSANFLQHL LHRHKFSYDTFVDYSIDEEAAFQAALALSLSEN
3676	A	3	1856	TLGRWLLGVYETVAPTLACLPRPRLRRRRRRR RRMISRYTRKAVPQSLELKGITKHALNHHPPPEK LEEISPTSDSHEKDTSSQSKSDITRESSFTSADTGN SLSAFPSYTGAGISTEGSSDFSWGYGELDQNATE KVQTMFTAIDELLYEQKLSVHTKSLQEECQQWT ASFPHLRILGRQIITPSEGYRLYPRSPSAVSASYET TLSQERDSTIFGIRGKKLHFSSSYAHKASSIAKSSS FCSMERDEEDSIIVSEGIIEEYLAFDHIDIEEGFHG KKSEAATEKQKLGYPPIAPFYCMKEDVLAYVFD SVWCKVVSCMEQLTRSHWEGFASDDESNVAVT RPDSESSCVLSELHPLVLPRVPQSKVLYITSNPMS LCQASRHQPNVNDLLVHGMPLQPRNLSLMDKLL DLDDKLLMRPGSSTILSTRNWPNRAVEFSTSSLS YTVQSTRRRNPPPRTLHPISTSHSCAETPRSVEEIL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RGARVPVAPDSLSSPSPTPLSRNNLLPPIGTAEVEHVSTVGPQRQMKPHGDSSRAQSAVVDEPNYQQPQERLLLPDFFPRPNTTQSFLLDTQYRRSCAVEYPHQARPGRGSAGPQLHGSTKSQSGGRPVSRTRQGP
3677	A	246	757	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWCEGPRMLSWCPFYKVLLLVQTAIYSVVGYASYLVWKDLGGGLGWPLALPLGLYAVQLTISWTVLVLFFTVHNPGLALLHLLLLYGLVVSTALIWHPINKLAALLLLPYLAWLTVTSALTYHLWRDSLCPVHQPQPTEKSD
3678	A	20	1508	RGKAEFFLAMAGTNALLMLENFIDGKFLPCSSYIDSYDPSTGEVYCRVPNSGKDEIEAAVKAAREAFPSWSSRSPQERSRVLNQVADLLEQSLEEFAQAESKDQGKTLALARTMDIPRSVQNFRFFASSSLHHTSECTQMDHLGCMHYTVRAPVGVAGLISPWNLPLYLLTWKIAPAMAAGNTVIAKPSELTSVTAWMLCKLLDKAGVPPGVVNIVFGTGPRVGEALVSHPEVPLISFTGSQPTAERITQLSAPHCKKLSLELGGKNPAIIFEDANLDECIPATVRSSFANQGEICLCTSRIFVQKSIYSEFLKRFVEATRKWKVGIPSDPLVSIGALISKAHLEKVRSYVKRALAEGAQIWCGEGVDKLSLPARNQAGYFMLPTVITDIKDESCCMTEEIFGPVTCVVPFDSEEEVIERANNVKYGLAATVWSSNVGRVHRVAKKLQSGLVWTNCWLIRELNLPFGGMKSSGIGREGAKDSYDFFTEIKTITVKH
3679	A	1862	502	MAGTKPYMEIQTTIREYYEHLYANKLENLEEMDKFLDTYTLPRLNQEEVESLNRPITGSEIEAIINSLPTKKIPGPDRFTAKFYQRYKEELSNLIHYLGLSHHLLALNFIIVSFGKKSAWSSAQVKVTDTDFDGVEVRVFEGPPKPEEPLKRSVVYIHGGGWALASAKIRYYDELCTAMAEELNAVIVSIEYRLVPKVYFPEQIHDVVRATKYFLKPEVLQKYMVDPGRICISGDSAGGNLAAALGQQFTQDASLKNKLKLQALIYPVLQALDFNTPSYQQNVNTPILPRYVMVKYWVDYFKGNYDFVQAMIVNNHTSLDVEEAAAVRARLNWTSLLPASFTKNYKPVVQTTGNARIVQELPQLLDARSAPLIADQAVLQLLPKTYILTCEHDVLRDDGIMYAKRLESAGVEVTLDHFEDGFHGCMIFTSWPTNFSVGIRTRNSYIKWLDQNL
3680	A	249	2146	RSWGAPWFWRMRLLRRRHMPLRLAMVGCAFVLFLFLLHRDVSSREEATEKPWLKSLVSRKDHVLDLMLEAMNNLRDSMPKLQIRAPEAQQTLFSINQSCLPGFYTPAELKPFWERPPQDPNAPGADGKAFQKSKWTPLETQEKEEGYKKHCFNAFASDRISLQRSLGPDTRPPECVDQKFRRCPPLATTSVIIVFHNEAWSTLLRTVYSVLHTTPAILLKEIILVDDASTEEHLKEKLEQYVKQLQVVRVVRQEERKGLITARLLGASVAQAEVLTFLDAHCECFHGWLEPLLARIAEDKTVVVSPDIVTIDLNTFEFAKPVQRGRVHSRGNFDWSLTFGWETLPPHEKQRRKDETYPIKSPTFAGGLFSISKSYFEHIGTYDNQMEIWGGENVEMSFRVWQCGGQLEIIPCSVVGHVFRTKSPHTFPKGTSVIARNQVRLAEVWMDSYKKIFYRRNLQAAKMAQEKSFGDISERLQLREQLHCHNFSWYLHNVYPEMFVPDL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TPTFYGAIKNLGTNQCLDVGENNRGGKPLIMYSCHGLGGNQYFEYTTQRDLRHNIAKQLCLHVSKGALGLGSCHFTGKNSQVPKDEEWELAQDQLIRNSGSGTCLTSQDKKPAMAPCNPSDPHQLWLFV
3681	A	2982	1869	LKDTLKSQMTQEASDEAEDMKEAMNRMIDELNKQVSELSQLYKEAQAELEDYRKRKSLEDVTAEYIHKAEHEKLMQLTNVSRAKAEDALSEMKSQYSKVLNELTQLKQLVDAQKENSVSITEHLQVITTLRTAAKEMEEKISNLKEHLASKEVEVAKLEKQLLEEKAAMTDAMVPRSSYEKLQSSLESEVSVLASKLKESVKEKEKVHSEVVQIRSEVSQVKREKENIQTLLKSKEQEVNELLQKFQQAQEELAEMKRYSESSSKLEEDKDKKINEMSKEVTKLKEALNSLSQLSYSTSSSKRQSQQLEALQQQVKQLQNQLAECKKQHQEVISVYRMHLLYAVQGQMDEDVQKVLKQILTMCKNQSQKK
3682	A	447	1024	AQALTAGRQLALAAPFIAPISPISLPRLNPPSQSWNSTPFFKVKLPPQKEVITSDELMAHLGNCLLSIKPQEKSEGLQLNFQQNVDDAMTVLPKLATGLDVNVRFTGVSDFEYTPECSVFDLLGIPLYHGWLVDPQQSPEAVRAVGKLSYNQL/VGEDHHLQTLQ*HQPRDRKPDCRAVPGDHRGPSDLPRTV
3683	A	2	942	LEIKQEEKFVGQCIKEELMHGECVKEEKDFLKKEIVDDTKVKEEPPINHPVGCKRKLAMSRCETCGTEEAKYRCPRCMRYSCSLPCVKKHKAELTCNGVRDKTAYISIQQFTEMNLLSDYRFLEDVARTADHISRDAFLKRPISNKYMYFMKNRARRQGINLKLLPNGFTKRKENSTFFDKKKQQFCWHVKLQFPQSQA\ST*KKRVPDDKTINEILKPYIDPEKSDPVIRQRLKAYIRSQTGVQILMKIEYMQQNLVRYYELDPYKSLLDNLRNKVIIEYPTLHVVLKGSNNDMKVLHQVKSESTKNVGNEN
3684	A	119	1533	SLQENVQEKRVRVCPGLGGLLPNGTPSITAAAAPQVLWRHVQPGCSHHLHACVIRAACRAGEGHADRHAGPPET/PVTLPSSWPWSSPWERQCPMH\L*APGHAFRPVPTEHRRGWAALGHHRAAAGPLREPASGSQPAPASC*PECHHGCPEQTRQCQDLLREAVVAPEQRG*PCAHLQT*ATATTLCPQVPAGRVWQPGHSCHLLPHRHDGSH*HHCAAHRRPVTRRQAAHGVPLPDACYSPHHTLPAAPPPATRPAGHTATHPE*GGDLTPVPDGPHDCPRDVQGIPGAGGGSQLAPCCPPFPAAPVSVQGTQGLGPKNVLH*QWEGIRWQKEPE/PGPPPEVELKRGAKCRIGDHGLGAVLGQGEYAS*SPSIPW*ASSSACPPLHPTP/TVYTQSPAAAPGWTRPPSP/PPPGLYPGP/PASHAPGVRGGISHQLYSLP*LCRECCSCP/PPPPAHGGRCPSLLPPEALAKLLL
3685	A	101	438	AWVLQCKINTELQTEVVMLKSMVLWLGEQVQSLQLQQQLHCHFNHTHICVTNLEYN\KEYPWDLVKAHLQGASTSNITFDIGELQKK\ILDLNKQTQEFQPSL*AWTEFQQGLE
3686	A	105	845	VSDVVKNQLVEVQCRQDGCDAVENVHQMFMFNWFTDCLWTLFLSNYQPSVESSSPGGSATSDDHEFDPSADMLVHDFDDERTLEEEEMMEGETNFSSEIEDLAREGDMPIHELLSLYGYGSTVRLPEEDEEEE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EEEEEGEDDEDADNDDNSGCSGENKEENIKDSSGQEDETQSSNDDPSQSVASQDAQEIIRPRRCKYFDTNSEVEEESEEDEDYIP/SIISFFQSSDGI*SSSSSEDWKKEIMVGS
3687	A	49	1225	PVLVTSLRMREADTLRPPQLMEVSADIISTVEFNHTGELLATGDKGGRVVIFQREPESKNAPHSQGEYDVYSTFQSHEPEFDYLKSLEIEEKINKIKWLPQQNAAHSLLSTNDKTIKLWKITERDKRPEGYNLKDEEGKLKDLSTVTSLQVPVLKPMDLMVEVSPRRIFANGHTYHINSISVNSDCETYMSADDLRINLWHLAITDRSFTP\NIVDIKPANMEDLTEVITASEFHPHHCNLFVYSSSKGSLRLCDMRAAALCDKHSKLFEEPEDPSNRSFFSEIIS\SVSDVKFSHSDRYMLTR\DYLTVKVWDLNMEARPIETYQVHDYLRSKLCSLYENDCIFDKFECAWNGSDR/IIMTGAYNNFFRMFDRNTKRDVTLEASRGSSKPRAVL
3688	A	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGPSLGLAAGIPLLVATALLVALLFTLIHRRRSSIEAMEESDRPCEISEIDDNPKISENPRRSPTHEKNTMGAQEAHIYVKTVAGSEEPVHDRYRPTIEMERRR
3689	A	698	889	GRVLVHCAMGVSRSATLVLAFLMIYENMTLVEAIPDGAGPPQISALTQAFVRQLQVLDNRLGRE
3690	A	61	153	MGAHLVRRYLGDASVEPDPLQMPTFPPDYGF
3691	A	61	153	MGAHLVRRYLGDASVEPDPLQMPTFPPDYGF
3692	A	3	2831	PLVRRLLRQTLRRVGGARAVREAVMRAVLTWRDKAEHCINDIAFKPDGTQLILAAGSRLLVYDTSDGTLLQPLKGHKDTVYCVAYAKDGKRFASGSADKSVIIWTSKLEGILKYTHNDAIQCVSYNPITHQLASCSSSDFGLWSPEQKSVSKHKSSSKIICCSWTNDGQYLALGMFNGIISIRNKNGEEKVKIERPGGSLSPIWSICWNPSSRWESFWMNRENEDAEDVIVNRYIQEIPSTLKSAVYSSQGSEAEEEEPEEEDDSPRDDNLEERNDILAVADWG\QKVSFYQLSGKQIGKDRALNFDPCCISYFTKGEYILLGGSDKQVSLFTKDGVRLGTVGEQNSWVWTGQAKPDSNYVVGGCQDGTISFYQLIFSTVHGLYKDRYAYRDSMTDVIVQHLITEQKVRIKCKELVKKIAIYRNRLAIQLPEKILIYELYSEDLSDMHYRVKEKIIKKFECNLLVVCANHIILCQEKRLQCLSFSGVKEREWQMESLIRYIKVIGGPPGREGLLVGLKNGQILKIFVDNLFAIVLLKQATAVRCLDMSASRKKLAVVDENDTCLVYDIDTKELLFQEPNANSVAWNTQCEDMLCFSGGGYLNIKASTFPVHRQKLQGFVVGYNGSKIFCLHVFSISAVEVPQSAPMYQYLDRKLFKEAYQIACLGVTDTDWRELAMEALEGLDFETAKKERKKRGETNNDLFLADVFSYQGKFHEAAKLYKRSGHENLALEMYTDLCMFEYAKDFLGSGDPKETKMLITKQADWARNIKEPKAAVEMYISAGEHVKAIEICGDHGWVDMLIDIARKLDKAEREPLLLCATYLKKLDSPGYAAETYLKMGDLKSLVQLHVETQRWDEAFALGEKHPEFKDDIYMPYAQWLAENDRFEEAQKAFHKAGRQREAVQVLEQLTNNAVAESRFNDAAYYYWMLSMQCLDIAQDPAQKD
3693	A	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTIHGGWRHHRDHTAIDEWDFNPSKFLIYTCLLLFSVLL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PLRLDGIIQWSYWAVFAPIWLWKLLVVAGASVGAGVWARNPRYRTEGEACVEFKAMLIAVGIHLLLLMFEVLVCDRVERGTHFWLLVFMPLFFVSPVSVAACVWGFRHDRSLELEILCSVNILQFIFIALKLDRIIHWPWLVVFVPLWILMSFLCLVVLYYIVWSLLFLRSLDVVAEQRRTHVTMAISWITIVVPLLTFEVLLVHRLDGHNTFSYVSIFVPLWLSLLTLMATTFRRKGGNHWWFAIRRDF/CQDQLPQPTGKPPPPPLTDHHGEKALPLQNKDRGSWPASRGSPRLL
3694	A	483	761	PRSLIDYKSYMDTKLLVARFLEQSSCTMTPDIHELVENIKSVLKSDEEHMEEAITSASFLEQIMAHSXQHIRAHKLPXETAGLXTSELRXLTP
3695	A	483	761	PRSLIDYKSYMDTKLLVARFLEQSSCTMTPDIHELVENIKSVLKSDEEHMEEAITSASFLEQIMAHSXQHIRAHKLPXETAGLXTSELRXLTP
3696	A	456	733	LSAALWEEPILSLWSETKELTNRGKMNYPQIGPHRPHVKGLRVRPGPGTLSNAPKSLCPGMSNSDRGIH\GGEGQGPGKRAGHLGRGGGMSFL
3697	A	877	1873	VWL*TLS*HTCALMTVCRSCLVKYLEENNTCPTCRIVIHQSHPLQYIGHDRTMQDIVYKLVPGLQEAEMRKQREFYHKLGMEVPGDIKGETCSAKQHLDSHRNGETKADDSSNKEAAE
3698	A	1	572	KQCGIPHEVVRDENSSVYAEVSRLLLATGHWKRLRRDNPRFNLMLGERNRLPFGRLGHEPGLVQLVNYYRGADKLCRKASLVKLIKTSPELAESCTWFPESYVIYPTNLKTPVAPAQNGIQPPISNSRTDEREFFLASYNRKKEDGEGNVWIAKSSAGAKVWVQW*MTDLEEEIDIPSPVGLGLESEWPL
3699	A	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKVEEHHLQPVQVLQTLLHSATAGTGCRRPARPPPAPPTPTPWRSRQSGKQSERAS*LKGRGRYGLGALGGRGGRALGGSRWPPPLPGETLFSGCKHRRRRRGSDAAPGEEAGT
3700	A	33	1318	GYQIGMALASGPARRALAGSGQLGLGGFGAPRRGAYEWGVRSTRKSEPPPLDRVYEIPGLEPITFAGKMHFVPWLARPIFPPWDRGYKDPRFYRSPPLHEHPLYKDQACYIFHHRCRLLEGVKQALWLTKTKLIEGLPEKVLSLVDDPRNHIENQDECVLNVISHARLWQTTEEIPKRETYCPVIVDNLIQLCKSQILKHPSLARRICVQNSTFSATWNRESLLLQVRGSGGARLSTKDPLPTIASREEIEATKNHVLETFYPISPIIDLHECNIYDVKNDTGFQEGYPYPYPHTLYLLDKANLRPHRLQPDQLRAKMILFAFGSALAQARLLYGNDAKVLEQPVVVQSVGTDGRVFHFLVFQLNTTDLDSNEGVKNLAWVDSDQLLYQHFWCLPVIKKRVVVEPVGPVGFKPETFRKFLALYLHGAA
3701	A	86	465	WTLCGPEAGMVGYDPKPDGRNNTKFQVAVAGSVSGLVTRALISPFDVIKIRFQLQHERLSRSDPSAKYHGILQASRQILQEEGPTAFWKGHVPAQILSIGYGAVQFLSFEMLTELVHRGSVYDARE
3702	A	166	814	GFWEKTNQSSHSMDPLGAPSQFVDVDTLPSWGDSCQDELNSSDTTAEIFQEDTVRSPFLYNKDVNGKVVLWKGDVALLNCTAIVNTSNESLTDKNPVSESIFMLAGPDLKEDLQKLKGCRTGEAQLTKGFNLAARFIIHTVGPKYKSRYRTAAESSLYSCYRNVLQLA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KEQSMSSVGFCVINSAKRGYPLKDATHIALRTVR RFLEIHGETIEKVV
3703	A	128	1255	SLGPSPKSATIPCCGDTMAPEEDAGGEALGGSFW EAGNYRRTVQRVEDGHRLCGDLVSCFQERARIE KAYAQQLADWARKWRGTVEKGPQYGTLEKAW HAFFTAAERLSALHLEVREKLQGQDSERVRAWQ RGAFHRPVLGGFRESRAAEDGFRKAQKPWLKRL KEVEASKKSYHAARKDEKTAQTRESHAKADSA VSQEQLRKLQERVERCAKEAEKTKAQYEQTLAE LHRYTPRYMEDMEQAFETCQAAERQRLLFFKD MLLTLHQHLDLSSSEKFHELHRDLHQGIEAASDE EDLRWWRSTHGPGMAMNWPQFEEWSLDTQRTI SRKEKGGRSPDEVTLTSIVPTRDGTAPPPQSPGSP GTGQDEEWSDEESP
3704	A	1	271	ARGEDLALATGGGPDTVTHSNMPCPNSLVYDC WLNIKECSVGEHTFEDLGLCPGRNQREKKRSYK DFLREEEKIAAQVRNSSKKKLKDSE
3705	A	170	1318	LNWANLVIMWPREEEKEKVQDYSLGGLSPDLRI DVSRKKKILKAYDEDEDEDLYPDIHPPPSLPLPG QFTCPQCRKSFTRRSFRPNLQLANMVQIIRQMCP TPYRGNRSNDQGMCFKHQEALKLFCEVDKEAIC VVCRESRSHKQHSVLPLEEVVQEYKAKLQGHVE PLRKHLEAVQKMKAKEERRVTELKSQMKSELA AVASEFGRLTRFLAEEQAGLERRLREMHEAQLG RAGAAASRLAEQAAQLSRLLAEAQERSQQGGLR LLQDIKETFNRCEEVQLQPPEVWSPDPCQPHSHD FLTDAIVRKMSRMFCQAARVDLTLDPDTAHPAL MLSPDRRGVRLAERRQEVADHPKRFSADCCVLG AQGFRSGRHYWEVCMGP
3706	A	204	1996	SRERQTTWMDHNFAPAPPEMQSHGAPGPGTSFS HSHVLGRPIRPSRLPGGGSPLTPVLRKTIHLDTFP QSHIPQTSSRLGLGARTRSVPPQETGIALGASLSP LPTSSLVPRKLSSISLTLHQNSQARSLDRPLSHWE ELPTPGKKAAPHEGGRVSSPGSPPVTLVPGGRVH SEGPGNPGLTKSNRMLATEKPLVSSYLALPFQSR LAQSAPVLAEPGSLGQGHLVSVTDHMPTRASPG KGKPRARGIPRPRGRLQRANTTVNLTAMDTRTD AARHLATMATNRPSLAINLATPNTSQLDTGTEFP ALDIKLGTARDLSSVGTVKSGKTVNLATAGTIKP GTAMNLTTVGTTKPGMVMDLIASEPDKLGKAM ATRSTAKPDMTTEGIAMDSATSDPVKPDTITATV GTSRLETAMALARVNRAKLGTAKNSLALDTSR MGTAVGSVVPVTPDPATGKTTLGSVNNLTISDV ATCLLMPSRSTDLALDNTNAAMDRATEPASLDL ATEYKGKCRNLVGDGLGCREGEVCELGDGSMK PMSINSNLLGYIGIDTIIEQMRKKTMKTGFDFNIM VVGTEGCGAAAGLVAGSTKDPISFPQ
3707	A	3	549	SSSISRDFLGQAACASGTMLRWLRDFVLPTAACQ DAEQPMRYETLFQALDRNGDGVVDIGELQEGLR NLGIPLGQDAEEKIFTTGDVNKDGKLDFEEFMKY LKDHEKKMKLAFKSLDKNNDGKIEASEIVQSLQ TLGLTISEQQAELILQSIDVDGTMTVDWNEWRD YFLFNPVTDIEEIIR
3708	A	1	1866	EFRGAGRANMLAPRGAAVLLLHLVLQRWLAAG AQATPQVFDLLPSSSQRLNPGALLPVLTDPALND

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LYVISTFKLQTKSSATIFGLYSSTDNSKYFEFTVM GRLSKAILRYLKNDGKVHLVVFNNLQLADGRRH RILLRLSNLQRGAGSLELYLDCIQVDSVHNLPRA FAGPSQKPETIELRTFQRKPQDFLEELKLVVRGSL FQVASLQDCFLQQSEPLAATGTGDFNRQFLGQM TQLNQLLGEVKDLLRQEVNETSFLRNTITECQAC GPLKFQSPTPSTVVPPASPAPPTRPPRRCDSNPCF RGVQCTDSRDGFQCGPCPEGYTGNGITCIDVDEC KYHPCYPGEHCINLSPGFRCDACPVGFTGPMVQ GVGISFAKSNKQVCTDIDECRNGACVPNSICVNT LGSYRCGPCKPGYTGDQIRGCKAERNCRNPELN PCSVNAQCIEERQGDVTCVCGVGWAGDGYICGK DVDIDSYPDEELPCSARNCKKDNCKYVPNSGQE DADRDGIGDACDEDADGDGILNEQDNCVLIHNV DQRNSDKDIFGDACDNCLSVLNNDQKDTDGDG RGDACDDDMDGDGIKNILDNCPKFPNRDQRDK DGDGVGDACDSCPDVSNPNQ
3709	A	144	417	TQAMEGLLHYINPAHAISLLSALNEERLKGQLCD VLLIVGDQKFRAHKNVLAASSEYFQSLFTNKENE SQTVFQLDFCEPDAFDNVLNYIY
3710	A	245	688	FGMLKNKGHSSKKDNLAVNAVALQDHILHDLQ LRNLSVADHSKTQVQKKENKSLKRDTKAIIDTGL KKTTQCPKLEDSEKEYVLDPKPPPLTLAQKLGLI GPPPPPLSSDEWEKVKQRSLLQGDSVQPCPICKE EFELRPQVFSIRG
3711	A	3	773	SLEMSSDGEPLSRMDSEDSISSTIMDVDSTISSGRS TPAMMNGQGSTTSSSKNIAYNCCWDQCQACFNS SPDLADHIRSIHVDGQRGGVFVCLWKGCKVYNT PSTSQSWLQRHMLTHSGDKPFKCVVGGCNASFA SQGGLARHVPTHFSQQNSSKVSSQPKAKEESPSK AGMNKRRKLKNKRRRSLARPHDFFDAQTLDAIR HRAICFNLSAHIESLGKGHSVVFHSTVSILLFFQIK YKTLQKNISTIISKSLKI
3712	A	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHSFLR GLFGGNTRIEEACEMYTRAANMFKMAKNWSAA GNAFCQAAKLHMQLQSKHDSATSFVDAGNAYK KADPQGKTARHVACYLCV
3713	A	20	974	GAAATACSSSSSSSGAPATWAAHGPGKDVASPS SVSLSPRRSRLLVLRCGLRRNPERPSSSPALRRLL LLLLLLLLLLGFLLSPGPERGVGGGRFGRRLAL LWAAALGHVVSGKVMSRRAPGSRLSSGGGGGG TNYSRSWNDWQPRTDSASADPGNLKYSSSRDRG GSSSYGLQPSNSAVVSRQRHDDTRVHADIQNDE KGGYSVNGGSGENTYGRKSLGQELRVNNVTSPE FTSVQHGSRALATKDMRKSQERSMSYCDESRLS YLLRRITRENDRDRRLATVKQLKEFIQQPENKLV LVKQLDILAAVHDVLNER
3714	A	237	458	IFALKSPSYLLPCCTPEGKMDHKQLCWSHPQKSG QSSRSCCICSNQHGLIWKYSLNMCLQCCHQYVK DIGFIKL
3715	A	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGFYH EAVVLFTQALKLNPQDHRLFGNRSFCHERLGQP AWALADAQVALTLRPGWPRGLFRLGKALMGLQ RFREAAAVFQETLRGGSQPDAARELRSCLLHLTL QGQRGGICAPPLSPGALQPLPHAELAPSGLPSLRC

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PRSTALRSPGLSPLLH
3716	A	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAMDSVPLISPLDISQLQPPLPDQVVIKTQTEYQLSSPDQQNYTKSR
3717	A	58	618	GAGCTSPGLWARKAAARCLPTYPSRAQPSNVGRRRRRRPGLGALAAGVPAMAESVERLQQRVQELERELAQERSLQVPRSGDGGGGRVRIEKMSSEVVDSNPYSRLMALKRMGIVSDYEKIRTFAVAIVGVGGVGSVTAEMLTRCGIGKLLLFDYDKVELANMNRLFFQPHQAGLSKVQAAGHTPEE
3718	A	3	593	RGAGGRAGGRADGQPNMADQRQRSLSTSGESLYHVLGLDKNATSDDIKKSYRKLALKYHPDKNPDNPEAADKFKEINNAHAILTDATKRNIYDKYGSLGLYVAEQFGEENVNTYFVLSSWWAKALFVFCGLLTCCYCCCCLCCCFNCCCGKCKPKAPEGEETEFYVSPEDLEAQLQSDEREATDTPIVIQPASATEP
3719	A	2	2173	SGGVRMGSRADGPRTSGHVTGKMAVFPWHSRNRNYKAEFASCRLEAVPLEFGDYHPLKPITVTESKTKKVNRKGSTSSTSSSSSSSVVDPLSSVLDGTDPLSMFAATADPAALAAAMDSSRRKRDRDDNSVVGSDFEPWTNKRGEILARYTTTEKLSINLFMGSEKGKAGTATLAMSEKVRTRLEELDDFEEGSQKELLNLTQQDYVNRIEELNQSLKDAWASDQKVKAPKNVHPGKLVYERIFSMCVDSRSVLPDHFSPENANDTAKETCLNWFFKIASIRELIPRFYVEASILKCNKFLSKTGISECLPRLTCMIRGIGDPL\GSVYARAYL\SRVGMEVAPHLKETLNKNFFDFLLTFKQIHGDTVQNQLVVQGVELPSYLPLYPPAMDWIFQCISYHAPEALLTEMMERCKKLGNNALLLNSVMSAFRAEFIATRSMDFIGMIKECDESGFPKHLLFRSLGLNLALADPPESDRLQILNEAWKVITKLKNPQDYINCAEVWVEYTCKHFTKREVNTVLADVIKHMTPDRAFEDSYPQLQLIIKKVIAHFHDFSVLFSVEKFLPFLDMFQKESVRVEVCKCI\RTPLSSINKSPPRTRSS*MPFCMFARPCMTL/CNALTLEDEKRMLSYLINGFIKMVSFGRDFEQQLSFYVESRSMFCNLEPVLVQLIHSVNRLAMETRKVMKGNHSRKTAAFVRSWGAYWFITIPSLAGIFTRLNLYLHSG
3720	A	24	296	ENLFRAGFAFSLLRSSFYISKTYCSWFSNLISGSLADFNSKGTRDYSPRQMAVRE/KVFDVIIRCFKRHGAEVIDTPVFELKVRNGQEETTW
3721	A	2	310	PSCLTCVGHCSIGGSCTMIGIMMPECHCSLHMTGPRCEEHVFILQQPGHIASILIPLLVLLLLALVAGVVFWHKRRVQGAKGFQHQRMTNGAMNVEIGNPTYK
3722	A	75	722	MELVAGCYEQVLFGFAVHPEPEACGDHEQWTLVADFTHHAHTASLSAVAVNSRFVVTGSKDETIHIYDMKKKIEHGALVHHSGTITCLKFYGNRHLISGAEDGLICIWDAKKWECLKSIKAHKGQVTFLSIHPSGKLALSVGTDKTLRTWNLVEGRSAFIKNIKQNAHIVEWSPRGEQYVVIIQNKIDIYQLDTASISGTITNEKRISSVKFLSES
3723	A	110	316	MELSDNRRSGGLEGLAEKCPNLTYLNLSGNKIKDLSTVEALVSGTVLSLDLLFLVKFSEICLCLLISI
3724	A	3	406	VDRGTEAWQRDPAFSGLQRVGGVDVSFVKGDS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VRACASLGVLSFPELEVVYEESRMVSLTAPYVSG FLAFREVPFLLELVQQLREKEPGLMPQVLLVDGN GVLHHRGFGVACHLGVLTDLPCVGVAKKLLQV DG
3725	A	3	406	VDRGTEAWQRDPAFSGLQRVGGVDVSFVKGDS VRACASLGVLSFPELEVVYEESRMVSLTAPYVSG FLAFREVPFLLELVQQLREKEPGLMPQVLLVDGN GVLHHRGFGVACHLGVLTDLPCVGVAKKLLQV DG
3726	A	1	433	SSDDRSLFRRLKLNYAIFDEGHMLKNMGSIRYQ HLMTINANNRLLLTGTPVQNNLLELMSLLNFVM PHMFSSSTSEIRRMFSSKTKSADEQSIYEKERIAH AKQIIKPFILRRVKEEVLKQLPPKKDRIELCAMSE KQEQLYLG
3727	A	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPCGW GQSSDLLSRIDLDELMKKDEPPLDFPDTLEGFEY AFNEKGQLRHIKTGEPFVFNYREHLHRWNQKRY EALGEIITKYVYELLEKDCNSKKVS
3728	A	3	2452	EIAGAAAENMLGSLLCLPGSGSVLLDPCTGSTISE TTSEAWSVEVLPSDSEAPDLKQEERLQELESCSG LGSTSDDTDVREVSSRPSTPGLSVVSGISATSEDIP NKIEDLRSECSSDFGGKDSVTSPDMDEITHDFLYI LQPKQHFQHIEAEADMRIQLSSSAHQLTSPPSQSE SLLAMFDPLSSHEGASAVVRPKVHYARPSHPPPD PPILEGAVGGNEARLPNFGSPMF*LPAEMEAFKQ RHS/YTPERLVRSRSS\DIVSSVRRPMSDPSWNRR P\GNEERELPPAAAIGATSLVAAPHSSSSSPSKDSS RGETEERKDSDDEKSDRNRPWWRKRFVSAMPK APIPFRKKEKQEKDKDDLGPDRFSTLTDDPSPRLS AQAQVAEDILDKYRNAIKRTSPSDGAMANYEST EVMGDGESAHDSPRDEALQNISADDLPDSASQA AHPQDSAFSYRDAKKKLRLALCSADSVAFPVLT\ HSTRNGLPDHTDPEDNEIVCFLKVQIAEAINLQD KNLMAQLQETMRCVCRFDNRTCRKLLASIAEDY RKRAPYIAYLTRCRQGLQTTQAHLERLLQRVLR DKEVANRYFTTVCVRLLLESKEKKIREFIQDFQK LTAADDKTAQVEDFLQFLYGAMAQDVIWQNAS EEQLQDAQLAIERSVMNRIFKLAFYPNQDGDILR DQVLHEHIQRLSKVVTANHRALQIPEVYLREAP WPSAQSEIRTISAYKTPRDKVQCILRMCSTIMNLL SLANEDSVPGADDFVPVLVFVLIKANPPCLLSTV QYISSFYASCLSGEESYWWMQFTAAVEFIKTIDD RK
3729	A	3	2452	EIAGAAAENMLGSLLCLPGSGSVLLDPCTGSTISE TTSEAWSVEVLPSDSEAPDLKQEERLQELESCSG LGSTSDDTDVREVSSRPSTPGLSVVSGISATSEDIP NKIEDLRSECSSDFGGKDSVTSPDMDEITHDFLYI LQPKQHFQHIEAEADMRIQLSSSAHQLTSPPSQSE SLLAMFDPLSSHEGASAVVRPKVHYARPSHPPPD PPILEGAVGGNEARLPNFGSPMF*LPAEMEAFKQ RHS/YTPERLVRSRSS\DIVSSVRRPMSDPSWNRR P\GNEERELPPAAAIGATSLVAAPHSSSSSPSKDSS RGETEERKDSDDEKSDRNRPWWRKRFVSAMPK APIPFRKKEKQEKDKDDLGPDRFSTLTDDPSPRLS AQAQVAEDILDKYRNAIKRTSPSDGAMANYEST

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EVMGDGESAHDSPRDEALQNISADDLPDSASQA AHPQDSAFSYRDAKKKLRLALCSADSVAFPVLT\ HSTRNGLPDHTDPEDNEIVCFLKVQIAEAINLQD KNLMAQLQETMRCVCRFDNRTCRKLLASIAEDY RKRAPYIAYLTRCRQGLQTTQAHLERLLQRVLR DKEVANRYFTTVCVRLLLESKEKKIREFIQDFQK LTAADDKTAQVEDFLQFLYGAMAQDVIWQNAS EEQLQDAQLAIERSVMNRIFKLAFYPNQDGDILR DQVLHEHIQRLSKVVTANHRALQIPEVYLREAP WPSAQSEIRTISAYKTPRDKVQCILRMCSTIMNLL SLANEDSVPGADDFVPVLVFVLIKANPPCLLSTV QYISSFYASCLSGEESYWWMQFTAAVEFIKTIDD RK
3730	A	3	2452	EIAGAAAENMLGSLLCLPGSGSVLLDPCTGSTISE TTSEAWSVEVLPSDSEAPDLKQEERLQELESCSG LGSTSDDTDVREVSSRPSTPGLSVVSGISATSEDIP NKIEDLRSECSSDFGGKDSVTSPDMDEITHDFLYI LQPKQHFQHIEAEADMRIQLSSSAHQLTSPPSQSE SLLAMFDPLSSHEGASAVVRPKVHYARPSHPPPD PPILEGAVGGNEARLPNFGSPMF*LPAEMEAFKQ RHS/YTPERLVRSRSS\DIVSSVRRPMSDPSWNRR P\GNEERELPPAAAIGATSLVAAPHSSSSSPSKDSS RGETEERKDSDDEKSDRNRPWWRKRFVSAMPK APIPFRKKEKQEKDKDDLGPDRFSTLTDDPSPRLS AQAQVAEDILDKYRNAIKRTSPSDGAMANYEST EVMGDGESAHDSPRDEALQNISADDLPDSASQA AHPQDSAFSYRDAKKKLRLALCSADSVAFPVLT\ HSTRNGLPDHTDPEDNEIVCFLKVQIAEAINLQD KNLMAQLQETMRCVCRFDNRTCRKLLASIAEDY RKRAPYIAYLTRCRQGLQTTQAHLERLLQRVLR DKEVANRYFTTVCVRLLLESKEKKIREFIQDFQK LTAADDKTAQVEDFLQFLYGAMAQDVIWQNAS EEQLQDAQLAIERSVMNRIFKLAFYPNQDGDILR DQVLHEHIQRLSKVVTANHRALQIPEVYLREAP WPSAQSEIRTISAYKTPRDKVQCILRMCSTIMNLL SLANEDSVPGADDFVPVLVFVLIKANPPCLLSTV QYISSFYASCLSGEESYWWMQFTAAVEFIKTIDD RK
3731	A	1	1305	VNTAMHEAKLMEECDELVEIIQQRKQMIAVKIK ETKVMKLRKLAQQVANCRQCLERSTVLINQAEH ILKENDQARFLQSAKNIAERVAMATASSQVLIPDI NFNDAFENFALDFSREKKLLEGLDYLTAPNPPSIR EELCTASHDTITVHWISDDEFSISSYELQYTIFTGQ ANFISLYNSVDSWMIVPNIKQNHYTVHGLQSGTR YIFIVKAINQAGSRNSEPTRLKTNSQPFKLDPKMT HKKLKISNDGLQMEKDESSLKKSHTPERFSGTGC YVYGVLHNSDNS*MFISLSFPLSHRYAIGIAYKSA PKNEWIGKNASSWVFSRCNSNFVVRHNNKEML VDVPPHLKRLGVLLDYDNY/NMLSFYDPANSL\H LHTFDVTF\ILPVCPTFTIWNKSLMILSGLPAPDFI DYPERQECNCRPQESPYVSGMKTCH
3732	A	127	2832	LGQRLSLVPRPSLKRRLGKRLSLGLRERMMSLW WS/GPKVRTQATTGARPKTETKSVPAARPKTEAQ AMSGARPKTEVQVMGGARPKTEAQGITGARPKT DARAVGGARSKTDAKAIPGARPKDEAQAWAQS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EFGTEAVSQAEGVSQTNAVAWPLATAESGSVTKSK\ACLWIEN*SMWM/PETFPGTQGQKGIQPWFGPGEETNMGSWCYSRPRAREEASNESGFWSADETSTASSFWTGEETSVRSWPREESNTRSRHRAKHQTNPRSRPRSKQEAYVDSWSGSEDEASNPFSFWVGENTNNLFRPRVREEANIRSKLRTNREDCFESESEDEFYKQSWVLPGEEAN\IDSGTETKKILILPWKLRAQKDVDSDRVKQEPRFEEEVIIGSWFWAEKEASLEGGASAICESEPGTEEGAIGGSAYWAEEKSSLGAVAREEAKPESEEEAIFGSWFWDRDEACFDLNPCPVYKVSDRFRDAAEELNASSRPQTWDEVTVEFKPGLFHGVGFRSTSPFGIPEEASEMLEAKPKNLELSPEGEEQESLLQPDQPSPEFTFQYDPSYRSVREIREHLRARESAESESWSCSCIQCELKIGSEEFEEFLLLMDKIRDPFIHEISKIAMGMRSASQFTRDFIRDSGVVSLIETLLNYPSSRVRTSFLENMIHMAPPYPNLNMIETFICQVCEETLAHSVDSLEQLTGNKGCFRHLTMTIDYHT\LIAN*YGPGFPLLF*PQAQCGETKFHVLKMLLNLSENPAVAKKLFSAKALSIFVGLFNIEETNDNIQIVIKMFQNISNIIKSGKMSLIDDDFSLEPLISAFREFEELAKQLQAQIDNQNDPEATGTTAFVGKGNNPSANRERLSPSVFCPGAQEAESLPARRVRGEEQRLLLEEVGARTADGIPEGW
3733	A	2	3274	DVPLIRIEEDTGEIFTTGARIDREKLCAGIPRDEHCFYEVEVAILPDEIFRLVKIRFLIEDINDNAPLFPATVINISIPENSAINSKYTLPAAVDPDVGINGVQNYELIKSQNIFGLDVIETPGGDKMPQLIVQKELDREEKDTYVMKVKVEDGGFPQRSSTAILQVSVTDTNDNHPVFKETEIEVSIPENAPVGTSVTQLHATDADIGENAKIHFSFSNLVSNIARRLFHLNATTGLITIKEPLDREETPNHKLLVLASDGGLMPARAMVLVNVTDVNDNVPSIDIRYIVNPVNDTVVLSENIPLNTKIALITVTDKDADHNGRVTCFTDHEIPFRLRPVFSNQFLLETAAYLDYESTKEYAIKLLA\ADAGKPPLNQSAMLFIKVKDENDNAPVFTQSFVTVSIPENNSPGIQLTKVSAMDADSGPNAKINYLLGPDAPPEFSLDCRTGMLTVVKKLDREKEDKYLFTILAKDNGVPPLTSNVTVFVSIIDQNDNSPVFTHNEYNFYVPENLPRHGTVGLITVTDPDYGDNSAVTLSILDENDDFTIDSQTGVIRPNISFDREKQESYTFYVKAEDGGRVSRSSSAKVTINVVDVNDNKPVFIVPPSNCSYELVLPSTNPGTVVFQVIAVDNDTGMNAEVRYSIVGGNTRDLFAIDQETGNITLMEKCDVTDLGLHRVLVKANDLGQPDSLFSVVIVNLFVNESVTNATLINELVPQKHLKHQ*PQILEIADVSSPTSDYVKILVAAVAGTITVVVVIFITAVVRCRQAPHLKAAQKNMQNSEWATPNPENRQMIMMKKKKKKKKHSPKNLLLNVVTIEETKADDVDSDGNRVTLDLPIDLEEQTMGKYNWVTTPTTFKPDSPDLARHYKSASPQPAFQIQPETPLNLKHHIIQELPLDNTFVACDSISNCSSSSSDPYSVSDCGYPVTTFEVPVSVHTRPPVDLEVGGAQSGQVAILTSSLMELLLCLMVAAFLPLELRPLGQQNVMSWEQEAKILLVGYWGDGEWCHFHFHHLIPGPVNPGYERKQYHILDSDSEDTQPSGELCPIPVRPFTILSIQLLQDDGEHCGTKQGFQPAVQLGLLPHKTLK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3734	A	1	840	GTRPGHLPAPSDGFCV/HL*SIPSWGSF*GESL/EMQLITSLGLQEFDIARNVLELIYAQTLVWIGIFFCPLLPFIQMIMLFIMFYSKNISLMMNFQPPSKAWRASQMMTFFIFLLFFPSFTGVLCTLAITIWRLKPSADCGPFRGLPLFIHSIYSWIDTLSTRPGYLWVVWIYRNLIGSVHFFFILTLIVLIITYLYWQITEGRKIMIRLLHEQIINEGKDKMFLIEKLIKLQDMEKKANPSSLVLERREVEQQGFLHLGEHDGSLDLRSRRSVQEGNPRA
3735	A	2	432	VEVCRRYLWKMTVDASQNVQCCVIFSHFPFIFNNLSKIKLLHTDTLLKIESKKHKAYLRSAAIEEERESEFALRPTFDLTVRRNHLIEDVLNQLSQFENEDLRKELWVSFSGEIGYDLGGS/VKKEIFYCLFAEMIQPEYGMFMY
3736	A	1542	343	KGAPSFVRLYQYPNFAGPHAALANKSFFKADKVTMLWNKKATAVLVIASTDVDKTGASYYGEQTLHYIATNGESAVVQLPKNGPIYDVVWNSSSTEFCAVYGFMPAKATIFNLKCDPVFDFGTGPRNAAYYSPHGHILVLAGFGNLILQI*AD/IMKVWNVKNYKLISKPVASDSTYFAWCPDGEHILTATCAPRLRVNNGYKIWHYTGSILHKYDVPSNAELWQVSWQPFLDGIFPAKTITYQAVPSEVPNEEPKVATAYRPPALRNKPITNSKLHEEEPPQNMKPQSGNDKPLSKTALKNQRKHEAKKAAKQEARSDKSPDLAPTPAPQSTPRNTVSQSISGDPEIDKKIKNLKKKLKAIEQLKEQAATGKQLEKNQLEKIQKETALLQELEDLELGI
3737	A	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPLKEEEILPEPGSETPTVASEALAELLHGALLRRGPEMGYLPGPPLGPEGGEEETTTTIITTTTVTTTVTSPVLCNNNISEGEGYVESPDLGSPVSRTLGLLDCTYSIHVYPGYGIEIQVQTLNLSQEEELLVLAGGGSPGLAPRLLANSSMLGEGQVLRSPTNRLLLHFQSPRVPRGGGFRIHYQAYLLSCGFPPRPAHGDVSVTDLHPGGTATFHCDSGYQLQGEETLICLNGTRPSWNGETPSCMASCGGTIHNATLGRIVSPEPGGAVGPNLTCRWVIEAAEGRRLHLHFERVSLDEDNDRLMVRSGGSPLSPVIYDSDMDDVPERGLISDAQSLYVELLSETPANPLLLSLRFEAFEEDRCFAPFLAHGNVTTTDPEYRPGALATFSCLPGYALEPPGPPNAIECVDPTEPHWNDTEPACKAMCGGELSEPAGVVLSPDWPQSYSPGQDCVWGVHVQEEKRILLQVEILNVREGDMLTLFDGDGPSARVLAQLRGPQPRRRLLSSGPDLTLQFQAPPGPPNPGLGQGFVLHFKEVPRNDTCPELPPPEWGWRTASHGDLIRGTVLTYQCEPGYELLGSDILTCQWDLSWSAAPPACQKIMTCADPGEIANGHRTASDAGFPVGSHVQYRCLPGYSLEGAAMLTCYSRDTGTPKWSDRVPKCALKYEPCLNPGVPENGYQTLYKHHYQAGESLRFFCYEGFELIGEVTITCVPGHPSQWTSQPPLCKVTQTTDPSRQLEGGNLALAILLPLGLVIVLGSGVYIYYTKLQGKSLFGFSGSHSYSPITVESDFSNPLYEAGDTREYEVSI
3738	A	3190	664	VAMGTPRAQHPPPPQLLFLILLSCPWIQGLPLKEEEILPEPGSETPTVASEALAELLHGALLRRGPEMGYLPGPPLGPEGGEEETTTTIITTTTVTTTVTSPVLCNNNISEGEGYVESPDLGSPVSRTLGLLDCTYSIHV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YPGYGIEIQVQTLNLSQEEELLVLAGGGSPGLAP RLLANSSMLGEGQVLRSPTNRLLLHFQSPRVPRG GGFRIHYQAYLLSCGFPPRPAHGDVSVTDLHPGG TATFHCDSGYQLQGEETLICLNGTRPSWNGETPS CMASCGGTIHNATLGRIVSPEPGGAVGPNLTCR WVIEAAEGRRLHLHFERVSLDEDNDRLMVRSGG SPLSPVIYDSDMDDVPERGLISDAQSLYVELLSET PANPLLLSLRFEAFEEDRCFAPFLAHGNVTTTDPE YRPGALATFSCLPGYALEPPGPPNAIECVDPTEPH WNDTEPACKAMCGGELSEPAGVVLSPDWPQSY SPGQDCVWGVHVQEEKRILLQVEILNVREGDML TLFDGDGPSARVLAQLRGPQPRRRLLSSGPDLTL QFQAPPGPPNPGLGQGFVLHFKEVPRNDTCPELP PPEWGWRTASHGDLIRGTVLTYQCEPGYELLGS DILTCQWDLSWSAAPPACQKIMTCADPGEIANG HRTASDAGFPVGSHVQYRCLPGYSLEGAAMLTC YSRDTGTPKWSDRVPKCALKYEPCLNPGVPENG YQTLYKHHYQAGESLRFFCYEGFELIGEVTITCV PGHPSQWTSQPPLCKVTQTTDPSRQLEGGNLAL AILLPLGLVIVLGSGVYIYYTKLQGKSLFGFSGSH SYSPITVESDFSNPLYEAGDTREYEVSI
3739	A	734	445	LLEPEPAEEYTEQSEVEST/EGMILI*CCLYFAAFQ TNVSNIYFALQYVNRQFMAETQFTSGEKEQVDE WTVETVEVRVLCIAKLLSLSSVSNFYLY
3740	A	2	1578	MAHYITFLCMVLVLLLQNSVLAEDGEVRSSCRT APTDLVFILDGSYSVGPENFEIVKKWLVNITKNF DIGPKFIQVGVVQYSDYPVLEIPLGSYDSGEHLTA AVESILYLGGNTKTGKAIQFALDYLFAKSSRFLT KIAVVLTDGKSQDDVKDAAQAARDSKITLFAIG VGSETEDAELRAIANKPSSTYVFYVEDYIAISKIR EVMKQKLCEESVCPTRIPVAARDERGFDILLGLD VNKKVKKRIQLSPKKIKGYEVTSKVDLSELTSNV FPEGLPPSYVFVSTQRFKVKKIWDLWRILTIDG/* PQIAVTLNGVDKILLFTTTSVINGSQVVTFANPQV KTLFDEGWHQIRLLVTEQDVTLYIDDQQIENKPL HPVLGILINGQTQIGKYSGKEETVQFDVQKLRIY CDPEQNNRETACEIPGFCLNGPSDVGSTPAPCICP PGKPGLQGPKGDPGLPGNPGYPGQPGQDGKPVS TESLVISGISGITGYQGIAGTPGVPGSPGIQGARGL. PGYKGEPGRDGDK
3741	A	5048	1236	MSAPAGSSHPAASARIPPKFGGSAVSGAAAPAGP GAGPAPHQQNGPAQNQMQVPSGYGLHHQNYIA PSGHYSQGPGKMTSLPLDTQCGDYYSALYTVPT QNVTPNTVNQQPGAQQLYSRGPPAPHIVGSTLGS FQGAASSASHLHTSASQPYSSFVNHYNSPAMYS ASSSVASQGFPSTCGHYAMSTVSNAAYPSVSYPS LPAGDTYGQMFTSQNAPTVRPVKDNSFSGQNTA ISHPSPLPPLPSQQHHQQQSLSGYSTLTWSSPGLP STQDNLIRNHTGSLAVANNNPTITVADSLSCPVM QNVQPPKSSPVVSTVLSGSSGSSSTRTPPTANHPV EPVTSVTQPSELLQQKGVQYGEYVNNQASSAPT PLSSTSDDEEEEEEDEEAGVDSSSTTSSASPMPNS YDALEGGSYPDMLSSSASSPAPDPAPEPDPASAP APASAPAPVVPQPSKMAKPLAMAIQHFSLVIRML QHHLFLEYSPSNPVYSGFQQYPQQYPGVNQLSSS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				IGGLSLQSSPQPESLRPVNLTQERNILPMTPVWAP VPNLNADLKKLNCSPDSFRCTLTNIPQTQALLNK AKLPLGLLLHPFRDLTQLPVITSNTIVRCRSCRTYI NP\FVSFIDQRR*KCNLCYRVNDVPEEFMYNPLT RSYGEPHKRPEVQNS\TVEFIASSDYMLRPPQPAV YLFVLDVSHNAVEAGYLTI/LWCQSLLE\NLDKLP G\DSRT\RIGFMTFD\STYSFLQFTQEGLSQPQMLI VSDIDDVFLPTPDSLLVNLYESKELIKDLLNALPN MFTNTRETHSALGPALQAAFKLMSPTGGRVSVF QTQLPSLGAGLLQSREDPNQRSSTKVVQHLGPAT DFYKKLALDCSGQQTAVDLFLLSSQYSDLASLA CMSKYSAGCIYYYPSFHYTHNPSQAEKLQKDLK RYLTRKIGFEAVMRIRCTKGLSMHTFHGNFFVRS TDLLSLANINPDAGFAVQLSIEESLTDTSLVCFQT ALLYTSSKGERRIRVHTLCLPVVSSLSDVYAGVD VQAAICLLANMAVDRSVSSSLSDARDALVNAVV DSLSAYGSTVSNLQHSALMAPSSLKLFPLYVLAL LKQKAFRTGTSTRLDDRVYAMCQIKSQPLVHLM KMIHPNLYRIDRLTDEGAVHVNDRIVPQPPLQKL SAEKLTREGAFLMDCGSVFYIWVGKGCDNNFIE DVLGYTNFASIPQKMTHLPELDTLSSERARSFIT WLRDSRPLSPILHIVKDESPAKAEFFQHLIEDRTE AAFSYYEFLLHVQQQICK
3742	A	934	68	SMLASQGVLLHPYGVPMIVPAAPYLPGLIQGNQE AAAAPDTMAQPYASAQFAPPQNGIPAEYTAPHP HPAPEYTGQTTVPEHTLNLYPPAQTHSEQSPADT SAQTVSGTRNKQD*RSTDGWPSPKTQTS*KHGK QVSSPSGLHVSNIPFR\FRDPDLRQMF\GQFGKILD VEIIFNERGSKGFGFVTFENSADADRAREK\LHGT VV\EGRKI\EVN\NATARVMTNKKTVNPYTNGWK LNPVVGAVYSPEFYAGTVLLCQANQEGSSMYSA PSTDFRGAKLHTSRPLLSGS
3743	A	3	1456	QFQQAWMQNKVPIPAPNEVLNDRKEDIKLEEKK KTQAEIEQEMATLQYTNPQLLEQLKIERLAQKQV EQIQPPPSSGTPLLGPQPFPGQGPMSQIPQGF/PTA PSISADANEHGS\KGPPGPQGQFRPPGPQGQMGP QGPPLHQGGGGPQGFMGPQGPQGPPQGLPRPQD MHGPQGMQRHPGPHGPLGPQGPPGPQGSSGPQG HMGPQGPPGPQGHIGPQGPPGPQGHLGPQGPPGT QGMQGPPGPRGMQGPPHPHGIQGGPGSQGIQGP VSQGPLMGLNPKGMQGPPGPRENQGPAPQGMI MGHPPQEMRGPHPPGGLLGHGPQEMRGPQEIRG MQGPPPQGSMLGPPQELRGPPGSQSQQGPPQGSL GPPPQGGMQGPPGPQGQQNPARGPHPSQGPIPFQ QQKTPLLGDGPRAPFNQEGQSTGPPPLIPGLGQQ GAQGRIPPLNPGQGPGPNKVS/ERGAPPRHEGRA PPRGRDGFPGPMKTLV
3744	A	1571	652	PLTGRKCPGWTHSGSRRSPRIAEEVPGFPKRAEA SRQFSETADRLELLRRAVMAAARATTPADGEEP APEAEALAAARERSSRFLSGLELVKQGAEARVFR GRFQGRAAVIKHRFPKGYRHPALEARLGRRRTV QEARALLRCRRAGISAPVVFFVDYASNCLYMEEI EGSVTVRD\IFSPLWRLKKTPQGLSNLAKTIGQVL ARMHDEDLIHGDLTTSNMLLKPPLEQLNIVLIDF GLSFISALPEDKGVDLYVLEKAFLSTHPNTETVFE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AFLKSYSTSSKKARPVLKKLDEVRLRGKKRSMVG
3745	A	127	1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLLWLALACSPVHTTLSKSDAKKAASKTLLEKSQFSDKPVQDRGLVVTDLKAESVVLEHRSYCSAKARDRHFAGDVLGYVTPWNSHGYDVTKVFGSKFTQISPVWLQLKRRGREMFEVTGLHDVDQGWMRAVRKHAKGL\P*CLGSCLRTGLTMISG/YVLDSEDEIEELSKTVVQVAKNQHFDGFVVEVWNQLLSQKRVGLIHMLTHLAEALHQARLLALLVIPPAITPGTDQLGMFTHKEFEQLAPVLDGFSLMTYDYSTAHQPGPNAPLSWVRACVQVLDPKSKWRSKILLGLNFYGMDYATSKDAREPVVGARYIQTLKDHRPRMVWDSQVSEHFFEYKKSRSGRHVVFYPTLKSLQVRLELARELGVGVSIWELGQGLDYFYDLL*VGIAASAVDVFFSKPWSE
3746	A	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLMVLYLIKKRLVACAAVFYGFAVHMKIYPETYILPITLHLLPDRDNDKSLRQFRYTFQACL*ELLKRLCNRTALMFVAVAGLTFFALSFGFYYEYGWEFLEHTYFYHLTRRDIRHNFSPYFYMLYLTAESKWSFSLGIAAFLPQLILLSAVSFAYYRDLVFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPLVMPLVRMPWKRAVVLLMLWFIGQAMWLAPAYVLEFQGKNTFLFIWLAGLFFLLINCSILIQIISHYKEEPLTERIKYD
3747	A	1	2325	MVISFQGLVTFGDVAVDFSQEEWEWLNPIQRNLYRKVMLENYRNLASLGLCVSKPDVISSLEQGKEPWTVKRKMTRAWCPDLKAVWKIKELPLKKDFCEGKLSQAVITERLTSYNLEYSLLGEHWDYDALFETQPGLVTIKNLAVDFRQQLHPAQKNFCKNGIWENNSDLGSAGHCVAKPDLVSLLEQEKEPWMVKRELTGSLFSGQRSVHETQELFPKQDSYAEGVTDRTSNTKLDCSSFRENWDSDYVFGRKLAVGQETQFRQEPITHNKTLSKERERTYNKSGRWFYLDDSEEKVHNRDSIKNFQKSSVVIKQTGIYAGKKLFKCNECKKTFTQSSSLTVHQRIHTGEKPYKCNECGKAFSDGSSFARHQRCHTGKKPYECIECGKAFIQNTSLIRHWRYYHTGEKPFDCIDCGKAFSDHIGLNQHRRIHTGEKPYKCDVCHKSF\RYGSSLTVHQRIHTGEKPYECDVCRKAFSHHASLT\Q\HQRVHSGEKPFKCKECGKAFRQNIHLASHLRIHTGEKPFECAECGKSFSISSQLATHQRIHTGEKPYECKVCSKAFTQKAHLAQHQKTHTGEKPYECKECGKAFSQTTHLIQHQRVHTGEKPYKCMECGKAFGDNSSCTQHQRLHTGQRPYECIECGKAFKTKSSLICHRRSHTGEKPYECSVCGKAFSHRQSLSVHQRIHSGKKPYECKECRKTFIQIGHLNQHKRVHTGERSYNYKKSRKVFRQTAHLAHHQRIHTGESSTCPSLPSTSNPVDLFPKFLWNPSSLPSP
3748	A	823	1	GGYTKSGYDSACKDFVPHDLEVQIPGRVFLVTGGNSGIGKATALEIAKRGGTVHLVCRDQAPAEDARGEIIRE\SGNQNIFLHIVDLSDPKKIWKFVENFKQEHKLHVL\VNNAGCMVNKREAHKKMDFEKNFGCQYSGVCTFLTTRPDPLCWRKNTDPRVIT\VSSGGMLVQKLNNQ*SPVRKNTIWMGTMVYAQNKVS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ERQQVVLT\ERWGPRAPG\IHFSSMHPGWA\DTPG VRQAMPGFHVQASGYRLRSEAQGADTMLWLAL SSARSRTAQRP
3749	A	1939	715	GFLRLSQAT\RQRLSIPVMVLTLDPTRD\QCFGDR FSRLLLDEFLGYDDIL\MSSVKGLAENEENKGFLR NVVSGEHYRFV\SMWMART\SYLAAFANHGQSF TLSVSHACCGYSHHQIFVFIVDLLQMLEMNMAIA FPAAPLLTVILALVGMEAIMSEFFNDTTTAFYIILI VWLADQYDAICCHTSTSKRHWLRFFYLYHFAFY AYHYRFNGQYSSLALVTSWLFIQHSMIYFFHHYE LPAILQHVRIQ\EMLLQAPTLGPGTPTA\LPDDMN NNSGAPATAP\DSAGQPPALGPVSPGASGSPGPV AAAPSSLVAAAASVAAAAGGDLGWMAETAAIIT DASFLSGLSASLLERRPASPLGPAGGLPHAPQDS VPPSDSAASDTTPLGAAVGGPSPASMAPTEAPSE VGS
3750	A	2	844	GLLEPFSKLLSFVIQNAVFTLAYLVELCGLCYRA FTKERDKFYLSRSVVLELLQALKLKSPLPDTNLL LLVQFICADAGTKLAESTILSKQMIASVPGCGTA AMECVRQYINEVLDFM\ADMHTLTKLKSHMKTC SQPLHEDTFGGHLKVGLAQIAAMDISRGNHRDN KAVIRYLPWLYHPPSAMQQGPKEFIECVSHIRLL SWLLLGSLTHNAVC/LKWPPLPGLPIPLDAGSHV ADHLIVILIGFPEQSKTSVL\HMCSLFHAF\SLAQL WDSLLARQSGRW
3751	A	431	2	AFTRKCEETAFIVPQCEIIPTE/WVCRRIPTGSSLER NPGVKEGCEFCPPKVEMFFKDDANHDPQWSRQ QLIAAKFGFAALGI/QTEVDIMSHAT*AVFEIPEKS RL\PQNCTPVDMKIEFGVHVTSKEILTDVIDNDS* RHSPS
3752	A	131	1278	AWSGSGLLVLCINTASMPMISVLGKMFLWQREG PGGRWTCQTSRRVSSDPAWAVEWIELPRGLSLSS LGSARTLRGWSRSSRPSSVDSQDLPEVNVGDTV AMLPKSRRALTIQEIAALARSSLHGISQVVKDHV TKPTAMAQGRVAHLIEWKGWSKPSDSPAALESA FSSYSDLSEGEQEARFAAGVAEQFAIAEAKLRA WSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPL GPHLQDLFTGHRFSRPVRQGSVEPESDCSQTVSP DTLCSSLCSLEDGLLGSPARLA\PSCWAMSCFSPN CPPAGKVPSAAW/APLEAQDSLYNSPLTESCLSP AEEEPAPCKDCQPLCPPLTGSWERQRQASDLASS GVVSLDEDEAEPEEQ
3753	A	3	1138	YYSSVRQRVTCEEPRFRECAAALIEGSATEVYAG EWRADRRSGFGVSQRSNGLRYEGEWLGNRRHG YGRTTRPDGSREEGKYKRNRLVHGGRVRSLLPL ALRRGKVKEKVDRAVEGARRAVSAARQRQEIA AARAADALLKAVAASSVAEKAVEAARMAKLIA QDLQPMLEAPGRRPRQDSEGSDTEPLDEDSPGV YENGLTPSEGSPELPSSPASSRQPWRPPACRSPLP PGGDQGPFSSPKAWPEEWGGAGAQAEELAGYE AEDEAGMQGPGPRDGSPLLGGCSDSSGSLREEE GEDEEPLPPLRAPAGTEPEPIAMLVLRGSSSRGPD AGCLTEELGEPAATERPAQPGAANPLVVGAVAL LDLSLAFLFSQLLT
3754	A	2	3338	SSLLEKMTSSDKDFRFMATSDLMSELQKDSIQLD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EDSERKVVKMLLRLLEDKNGEVQNLAVKWLGV PLGAFHASLLHCLLPQLSSPRLAVRKRAVGALGH LATACSTDLFVELADHLLDRLPGPRVPTSPTAIRT LIQCLGSVGRQAGHRLGAHLDRLVPLVEDFCNL DDDELRESCLQAFEAFLRKCPKEMGPHVPNVTS LCLQYIKHDPNYNYDSDEDEEQMETEDSEFSEQE SEDEYSDDDDMSWKVRRAAAKCIAALISSRPDL LPDFHCTLAPVLIRRFKEREENVKADVFTAYIVL LRQTRPPKGWLEAMEEPTQTGSNLHMLRGQVPL VVKALQRQLKDRSVRARQGCFSLLTELAGVLPG SLAEHMPVLVSGIIFSLADRSSSSTIRMDALAFLQ GLLGTEPAEAFHPHLPILLPPVMACVADSFYKIA AEALVVLQELVRALWPLHRPRMLDPEPYVGEMS AVTLARLRATDLDQEVKERAISCMGHLVGHLGD RLGDDLEPTLLLLLDRLRNEITRLPAIKALTLVAV SPLQLDLQPILAEALHILASFLRKNQRALRLATLA ALDALAQSQGLSLPPSAVQAVLAELPALVNESD MHVAQLAVDFLATVTQAQPASLVEVSGPVLSEL LRLLRSPLLPAGVLAAAEGFLQALVGTRPPCVDY AKLISLLTAPVYEQAVDGGPGLHKQVFHSLARC VAALSAACPQ\EAESTASRLVCDARSPHSSTGVK VLAFLSLAEVGQVAGPGHERELKAVLLEALGSPS EDVRAAASYALGRVGAGSLPDFLPFLLEQIEAEP RRQYLLLHSLKEALGAAQPDSLKPYAEDIWALL FQRCEGAEEGTRGVVAECIGKLVLVNPSFLLPRL RKQLAAGRPHTRSTVITAVKFLISDQPHPIDPLLK SFIAVHNKPSLVRDLLDDILPLLYQETKIRRDLIRE VEMGPFKHTVDDGLDVRKAAFECMYSLLESCLG QLDICEFLNHVEDGLKDHYDIRMLTFIMVARLAT LCPAPVLQRVDRLIEPLRATCTAKVKAGSVKQEF EKQDELKRSAMRAVAALLTIPEVGKSPIMADFSS QIRSNPELAALFESIQKDSTSAPSTDSMELS
3755	A	2	3338	SSLLEKMTSSDKDFRFMATSDLMSELQKDSIQLD EDSERKVVKMLLRLLEDKNGEVQNLAVKWLGV PLGAFHASLLHCLLPQLSSPRLAVRKRAVGALGH LATACSTDLFVELADHLLDRLPGPRVPTSPTAIRT LIQCLGSVGRQAGHRLGAHLDRLVPLVEDFCNL DDDELRESCLQAFEAFLRKCPKEMGPHVPNVTS LCLQYIKHDPNYNYDSDEDEEQMETEDSEFSEQE SEDEYSDDDDMSWKVRRAAAKCIAALISSRPDL LPDFHCTLAPVLIRRFKEREENVKADVFTAYIVL LRQTRPPKGWLEAMEEPTQTGSNLHMLRGQVPL VVKALQRQLKDRSVRARQGCFSLLTELAGVLPG SLAEHMPVLVSGIIFSLADRSSSSTIRMDALAFLQ GLLGTEPAEAFHPHLPILLPPVMACVADSFYKIA AEALVVLQELVRALWPLHRPRMLDPEPYVGEMS AVTLARLRATDLDQEVKERAISCMGHLVGHLGD RLGDDLEPTLLLLLDRLRNEITRLPAIKALTLVAV SPLQLDLQPILAEALHILASFLRKNQRALRLATLA ALDALAQSQGLSLPPSAVQAVLAELPALVNESD MHVAQLAVDFLATVTQAQPASLVEVSGPVLSEL LRLLRSPLLPAGVLAAAEGFLQALVGTRPPCVDY AKLISLLTAPVYEQAVDGGPGLHKQVFHSLARC VAALSAACPQ\EAESTASRLVCDARSPHSSTGVK VLAFLSLAEVGQVAGPGHERELKAVLLEALGSPS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EDVRAAASYALGRVGAGSLPDFLPFLLEQIEAPRRQYLLLHSLKEALGAAQPDSLKPYAEDIWALLFQRCEGAEEGTRGVVAECIGKLVLVNPSFLLPRLRKQLAAGRPHTRSTVITAVKFLISDQPHPIDPLLKSFIAVHNKPSLVRDLLDDILPLLYQETKIRRDLIREVEMGPFKHTVDDGLDVRKAAFECMYSLLESCLGQLDICEFLNHVEDGLKDHYDIRMLTFIMVARLATLCPAPVLQRVDRLIEPLRATCTAKVKAGSVKQEFEKQDELKRSAMRAVAALLTIPEVGKSPIMADFSSQIRSNPELAALFESIQKDSTSAPSTDSMELS
3756	A	112	1361	SLEEQQGRHPSFAPKCASQILGRIMITLITEQLQKQTLDELKCTRFSISLPLPDHADISNCGNSFQLVSEGASWRGLPHCSCAEFQ/DQPQLQLPSLRPEPAPQTT\HRGNSPKEQPFSQVLRPEPPDPEKLPVPPAPPSKRHCRSLSVPVDLSRWQPVWRPAPSKLWTPIKHRGSGGGGGPQVPHQSPPKRVSSL/SVPPSSQCLFSMCPSSHTLQPSFLQPGPGP\DSSRPCAASPQSGSWESDAESLSPCPPQRRFSLSPSLGPQASRFLPSARSSPASSPELPWRPRGLRNLPRSRSQPCDLDARKTGVKRRHEEDPRRLRPSLDFDKMNQKPYSGGLCLQETAREGSSISPPWFMACSPPPLSASCSPTGGSSQVLSESEEEEEGAVRWGRQALSKRTLCQRDFGDLDLNLIEEN
3757	A	413	1	PKPMLQQDFT/SLPDQGLDHIAE/NSYFDARSLCAAELVCKEWQQVTSE*MLWKKLIERMVHAYPLWKGLSEKVW/DQHLFKNRPTDGPPNSFHRSLYPKIIQVIETIESNWQCG*HTLQRIQCHSEKSKGVYCLQYDDEK
3758	A	2	613	FVSGSPWRMDGSTERLEARRPAGRLPWSSRQEMTRRPSLMAGRQHGWSAQQSATVANPVPGANPDLLPHFLGEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTMEVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSLEQGIVLPCRPPEGIPPAE
3759	A	1	561	ADDTLHLWNLRQKRPAILHSLKFCRERVTFCHLPFQSKWLYVGTERGNIHIVNVESFTLSGYVIMWNKAIELSSKSHPGPVVHISDNPMDEGKLLIGFESGTVVLWDLKSKKADYRYTYDEAIHSVAWHHEGKQFICSHSDGTLTIWNVRSPAKPVQTITPHGKQLKDGKKPEPCKPILKVEFXTTR
3760	A	1	824	LPACRCGCVAGCPSNHGICRCLRASERQVCVMHLKHLRTLLSPQDGAAKVTCMAWSQNNAKFAVCTVDRVVLLYDEHGERRDKFSTKPADMKYGRKSYMVKGMAFSPDSTKIAIGQTDNIIYVYKIGEDWGDKKVICNKFIQTVKFRPVPGTLG*TNIYQYIYL*IQPGVAFLTSECDFSYCKDGASWLFMVICCLP*SPAVSFPIGD*\SAVTCLQWPAEYIIVFGLAEGKVRLSNTKTNKSSTIYGTESYVVSLTTNCSGKGILSGHADGYQR
3761	A	2253	320	PVIQRCSQPYGFSLLISFFLKCVSETSQQPPSRKVFQLLPSFPTLTRSKSHESQLGNRIDDVSSMRFDLSHGSPQMVRRDIGLSVTHRFSTKSWLSQVCHVCQKSMIFGVKCKHCRLKCHNKCTKEAPACRISFLPLTRLRRTESVPSDINNPVDRAAEPHFGTLPKALTKK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EHPPAMNHLDSSSNPSSTTFSTPSSPAPFPTSSNPSSATTPP\NPSP\GQR\DSRFNFPSC/AYFIHHR\Q\QFIFPDISAFAHAAPLPEAADGTRLDDQPKADVLEAHEAEAEEPEAGKSEAEDDEDEVDDLPSSRRPWRGPISRKASQTSVYLQEWDIPFEQVELGEPIGQGRWGRVHRGRWHGEVAIRLLEMDGHNQDHLKLFKKEVMNYRQTRHENVVLFMGACMNPPHLAIITSFCKGRTLHSFVRDPKTSLDINKTRQIAQEIIKGMGYLHAKGIVHKDLKSRNVFYDNG\KVVITDFGLF\GISGVVP\EGRRENQLKLSHDWLCYLAPEIVREMTPGKDEDQLPFSKAADVYAFGTVWYELQARDWPLKNQAAEASIWQIGSGEGMKRVLTSVSLGKEVSENLSACWAFDLQERPS\FSLLMDMLEKLPKLNRRLSHPGHF*KSADINSSKVVPRFERFGLGVLESSNPKM
3762	A	2	1578	MAHYITFLCMVLVLLLQNSVLAEDGEVRSSCRTAPTDLVFILDGSYSVGPENFEIVKKWLVNITKNFDIGPKFIQVGVVQYSDYPVLEIPLGSYDSGEHLTAAVESILYLGGNTKTGKAIQFALDYLFAKSSRFLTKIAVVLTDGKSQDDVKDAAQAARDSKITLFAIGVGSETEDAELRAIANKPSSTYVFYVEDYIAISKIREVMKQKLCEESVCPTRIPVAARDERGFDILLGLDVNKKVKKRIQLSPKKIKGYEVTSKVDLSELTSNVFPEGLPPSYVFVSTQRFKVKKIWDLWRILTIDG/*PQIAVTLNGVDKILLFTTTSVINGSQVVTFANPQVKTLFDEGWHQIRLLVTEQDVTLYIDDQQIENKPLHPVLGILINGQTQIGKYSGKEETVQFDVQKLRIYCDPEQNNRETACEIPGFCLNGPSDVGSTPAPCICPPGKPGLQGPKGDPGLPGNPGYPGQPGQDGKPVSTESLVISGISGITGYQGIAGTPGVPGSPGIQGARGLPGYKGEPGRDGDK
3763	A	3	1267	CKVWRNPLNLFRGAEYNRYTWVTGREPLTYYDMNLSAQDHQTFFTCDSDHLRPADAIMQKAWRERNPQARISAAHEALEINECATAYILLAEEEATTIAEAEKLFKQALKAGDGCYRRSQQLQHHGSQYEAQHSVLYLPLQ\TRHQCLGVHQKKASNVCQKTREDQGSSENDERFNEGVPPSEYVQYP*KPF\KALLELQAYADVQAVLAKYDDISLPKSATICYTAALLKARAVSDKFSPEAASRRGLSTAEMNAVEAIHRAVEFNPHVPKYLLEMKSLILPPEHILKRGDSEAIAYAFFHLAHWKRVEGALNLLHCTWEGTFRMIPYPLEKGHLFYPYPICTETADRELLPSFHEVSVYPKKELPFFILFTAGLCSFTAMLALLTHQFPELMGVFAKAVSVCLEGGLGEWMGKAKGIKAA
3764	A	25	1032	RSADGLCGNKDRERGNEFTRNQQAAQEVVNPKKKMKKKKYVNSGTVTLLSFAVESECTFLDYIKGGTQINFTVAIDFTASNGNPSQSTSLHYMSPYQLNAYALALTAVGEIIQHYDSDKMFPALGFGAKLPPDGRVSHEFPLNGNQENPSCCGIDGILEAYHRSLRTVQLYGPTNFAPVVTHVARNAAAVQDGSQYSVLLIITDGVISDMAQTKEAIVNG\SKLPMSIIIVGVGQAEFNAMVELDGDDVRISSRGKLAERDIVQFVPFRDYVDRTGNHVLSMARLARDVLAEIPDQLVSYMKAQGIRPRSPPAAPTHSPSQSPARTPPACPLHTHI
3765	A	172	3456	LGMMDSPKIGNGLPVIGPGTDIGISSLHMVGYLG

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KNFDSAKVPSDEYCPACKEKGKLKALKTYRISFQ ESIFLCEDLQCIYPLGSKSLNNLISPDLEECHTPHK PQKRKSLESSYKDSLLLANSKKTRNYIAIDGGKV LNSKHNGEVYDETSSNLPDSSGQQNPIRTADSLE RNEILEADTVDMATTKDPATVDVSGTGRPSPQN EGCTSKLEMPLESKCTSFPQALCVQWKNAYALC WLDCILSALVHSEELKNTVTGLCSKEESIFWRLL TKYNQANTLLYTSQLSGVKDGDCKKLTSEIFAEI ETCLNEVRDEIFISLQPQLRCTLGDMESPVFAFPL LLKLETHIEKLFLYSFSWDFECSQCGHQYQNRH MKSLVTFTNVIPEWHPLNAAHFGPCNNCNSKSQI RKMVLEKVSPIFMLHFVEGLPQNDLQHYAFHFE GCLYQITSVIQYRANNHFITWILDADGSWLECDD LKGPCSERHKKFEVPASEIHIVIWERKISQVTDKE AACLPLKKTNDQHALSNEKPVSLTSCSVGDAAS AETASVTHPKDISVAPRTLSQDTAVTHGDHLLSG PKGLVDNILPLTLEETIQKTASVSQLNSEAFL\LEN KPVAENTGILKTNTLLSQESLMASSVSAPCNEKLI QDQFVDISFPSQVVNTNMQSVQLNTEDTVNTKS VNNTDATGLIQGVKSVEIEKDAQLKQFLTPKTEQ LKPERVTSQVSNLKKKETTADSQTTTSKSLQNQS LKENQKKPFVGSWVKGLISRGASFMPLCVSAHN RNTITDLQPSVKGVNNFGGFKTKGINQKASHVSK KARKSASKPPPISKPPAGPPSSNGTAAHPHAHAA SEVLEKSGSTSCGAQLNHSSYGNGISSANHEDLV EGQIHKLRLKLRKKLKAEKKKLAALMSSPQSRT VRSENLEQVPQDGSPNDCESIEDLLNELPYPIDIA NESACTTVPGVSLYSSQTHEEILAELLSPTPVSTE LSENGEGDFRYLGMGDSHIPPPVPSEFNDVSQNT HLRQDHNYCSPTKKNPCEVQPDSLTNNACVRTL NLESPMKTDIFDEFFSSSALNALANDTLDLPHFDE YLFENY
3766	A	3	1622	AQQIVYRNVMLENYKNLVSLGYQLTKPDVILRL EKGEEPWLVEREIHQETHPDSETAFEIKSSVSSRSI FKDKQSCDIKMEGMARNDLWYLSLEEVWKCRD QLDKYQENPERHLRQVAFTQKKVLTQERVSESG KYGGNCLLPAQLVLREYFHKRDSHTKSLKHDLV LNGHQDSCASNSNECGQTFCQNIHLIQFARTHTG DKSYKCPDNDNSLTHGSSLGISKGIHREKPYECK ECGKFFSWRSNLTRHQLIHTGEKPYECKECGKSF SRSSHLIGHQKTHTGEEPYECKECGKSFSWFSHL VTHQRTHTGDKLYTCNQCGKSF/VHSSRLIRHQR THTGEKPYECPECGKSFRQSTHLILHQRTHVRVR PYECNECGKSYSQRSHLVVHHRIHTGLKPFECKD CGKCFSRSSHLYSHQRTHTGEKPYECHDCGKSFS QSSALIVHQRIHTGEKPYECCQCGKAFIRKNDLIK HQRIHVGEETYKCNQCGIIFSQNSPFIVHQIAHTG EQFLTCNQCGTALVNTSNLIGYQTNHIRENAY
3767	A	3	1622	AQQIVYRNVMLENYKNLVSLGYQLTKPDVILRL EKGEEPWLVEREIHQETHPDSETAFEIKSSVSSRSI FKDKQSCDIKMEGMARNDLWYLSLEEVWKCRD QLDKYQENPERHLRQVAFTQKKVLTQERVSESG KYGGNCLLPAQLVLREYFHKRDSHTKSLKHDLV LNGHQDSCASNSNECGQTFCQNIHLIQFARTHTG DKSYKCPDNDNSLTHGSSLGISKGIHREKPYECK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ECGKFFSWRSNLTRHQLIHTGEKPYECKECGKSF SRSSHLIGHQKTHTGEEPYECKECGKSFSWFSHL VTHQRTHTGDKLYTCNQCGKSF/VHSSRLIRHQR THTGEKPYECPECGKSFRQSTHLILHQRTHVRVR PYECNECGKSYSQRSHLVVHHRIHTGLKPFECKD CGKCFSRSSHLYSHQRTHTGEKPYECHDCGKSFS QSSALIVHQRIHTGEKPYECCQCGKAFIRKNDLIK HQRIHVGEETYKCNQCGIIFSQNSPFIVHQIAHTG EQFLTCNQCGTALVNTSNLIGYQTNHIRENAY
3768	A	185	2258	SIIIKMSRKISKESKKVNISSSLESEDISLETTVPTD DISSSEEREGKVRITRQLIERKELLHNIQLLKIELS QKTMMIDNLKVDYLTKIEELEEKLNDALHQKQL LTLRLDNQLAFQQKDASKYQELMKQEMETILLR QKQLEETNLQLREKAGDVRRSLRDFELTEEQYIK LKAFPEDQLSIPEYVSVRFYELVNPLRKEICELQV KKNILAEELSTNKNQLKQLTETYEEDRKNYSEV QIRCQRLALELADTKQLIQQGDYRQENYDKVKS ERDALEQEVIELRRKHEILEASHMIQTKERSELSK EVVTLEQTVTLLQKDKEYLNRQNMELSVRCAHE EDRLERLQAQLEESKKAREEMYEKYVASRDHY KTEYENKLHDELEQIRLKTNQEIDQLRNASREMY ERENRNLREARDNAVAEKERAVMAEKDALEKH DQLLDRYRE\LQ\LSTESKVTEFLHQSKLKSFESE RVQLLQEETARNLTQCQLECEKYQKKLEVLTKE FYSLQASSEKRITELQAQNSEHQARLDIYEKLEK ELDEIIMQTAEIENEDEAERVLFSYGYGANVPTT AKRRLKQSVHLARRVLQLEKQNSLI/LKRSGTSK GPSNTAFTRSLTEANSLLNQTQQPYRYLIESVRQ RDSKIDSLTESIAQL/ERKDVSNLNKEKSALLQTN GIKMAL\DL\DQLLNHP
3769	A	3	2297	DAAEFRVVADAMKVIGFKPEEIQTVYKILAAILH LGNLKFVVDGDTPLIENGKVVSIIAELLSTKTDM VEKALLYRTVATGRDIIDKQHTEQEASYGRDAF AKAIYERLFCWIVTRINDIIEVKNYDTTIHGKNTV IGVLDIYGFEIFDNNSFEQFCINYCNEKLQQLFIQL VLKQEQEEYQREGIPWKHIDYFNNQIIVDLVEQQ HKGIIAILDDACMNVGKVTDEMFLEALNSKLGK HAHFSSRKLCASDKILEFDRDFRIRHYAGDVVYS VIGFIDKNKDTLFQDFKRLMYNSSNPVLKNMWP EGKLSITEVTKRPLTAATLFKNSMIALVDNLASK EPYYVRCIKPNDKKSPQIFDDERCRHQVEYLGLL ENVRVRRAGFAFRQTYEKFLHRYKMISEFTWPN HDLPSDKEAVKKLIERCGFQDDVAYGKTKIFIRT PRTLFTLEELRAQMLIRIVLFLQKVWRGTLARMR YKRTKAALTIIRYYRRYKVKSYIHEVARRFHGVK TMRDYGKHVKWPSPPKVLRRFEEALQTIFNRWR ASQLIKSIPASDLPQVRAKVAAVEMLKGQRADL GLQRAWEGNYLASKPDTPQTSGTFVPVANELKR KDKYMNVLFSCHVRKVNRFSKVEDRAIFVTDRH LYKMDPTKQYKVMKTIPLYNLTGLSVSNGKDQL VVFHTKDNKDLIVCLFSKQPTHESRIGEL\VGVLV NHFKSEKRHLQV\NVTNPVQCSLHGKKCTVSVE TRLNQPQPDFTKNRSGFILSVPGN
3770	A	3	6276	HKVAAPDVVVPTLDTVRHEALLYTWLAEHKPL VLCGPPGSGKTMTLFSALRALPDMEVVGLNFSS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ATTPELLLKTFDHYCEYRRTPNGVVLAPVQLGK WLVLFCDEINLPDMDKYGTQRVISFIRQMVEHG GFYRTSDQTWVKLERIQFVGACNPPTDPGRKPLS HRFLRHVPVVYVDYPGPASLTQIYGTFNRAMLR LIPSLRTYAEPLTAAMVEFYTMSQERFTQDTQPH YIYSPREMTRWVRGIFEALRPLETLPVEGLIRIWA HEALRLFQDRLVEDEERRWTDENIDTVALKHFP NIDREKAMSRPILYSNWLSKDYIPVDQEELRDYV KARLKVFYEEELDVPLVLFNEVLDHVLRIDRIFR QPQGHLLLIGVSGAGKTTLSRFVAWMNGLSVYQ IKVHRKYTGEDFDEDLRTVLRRSGCKNEKIAFIM DESNVLDSGFLERMNTLLANGEVPGLFEGDEYA TLMTQCKEGAQKEGLMLDSHEELYKWFTSQVIR NLHVVFTMNPSSEGLKDRAATSPALFNRCVLNW FGDWSTEALYQVGKEFTSKMDLEKPNYIVPDYM PVVYDKLPQPPSHREAIVNSCVFVHQTLHQANA RLAKRGGRTMAITPRHYLDFINHYANLFHEKRSE LEEQQMHLNVGLRKIKETVDQVEELRRDLRIKS QELEVKNAAANDKLKKMVKDQQEAEKKKVMS QEIQEQLHKQQEVIADKQMSVKEDLDKVEPAVI EAQNAVKSIKKQHLVEVRSMANPPAAVKLALES ICLLLGESTTDWKQIRSIIMRENFIPTIVNFSAEEIS DAIREKMKKNYMSNPSYNYEIVNRASLACGPMV KWAIAQLNYADMLKRVEPLRNELQKLEDDAKD NQQKANEVEQMIRDLEASIARYKEEYAVLISEAQ AIKADLAAVEAKVNRSTALLKSLSAERERWEKT SETFKNQMSTIAGDCLLSAAFIAYAGYFDQQMR QNLFTTWSHHLQQANIQFRTDIARTEYLSNADER LRWQASSLPADDLCTENAIMLKRFNRYPLIIDPS GQATEFIMNEYKDRKITRTSFLDDAFRKNLESAL RFGNPLLVQDVESYDPVLNPVLNREVRRTGGRV LITLGDQDIDLSPSFVIFLSTRDPTVEFPPDLCSRV TFVNFTVTRSSLQSQCLNEVLKAERPDVDEKRSD LLKLQGEFQLRLRQLEKSLLQALNEVKGRILDDD TIITTLENLKREAAEVTRKVEETDIVMQEVETVS QQYLPLSTACSSIYFTMESLKQIHFLYQYSLQFFL DIYHNVLYENPNLKGVTDHTQRLSIITKDLFQVA FNRVARGMLHQDHITFAMLLARIKLKGTVGEPT YDAEFQHFLRGNEIVLSAGSTPRIQGLTVEQAEA VVRLSCLPAFKDLIAKVQADEQFGIWLDSSSPEQ TVPYLWSEETPATPIGQAIHRLLLIQAFRPDRLLA MAHMFVSTNLGESFMSIMEQPLDLTQIVGTEVKP NTPVLMCSVPGYDASGHVEDLAAEQNTQITSIAI GSAEGFNQADKAINTAVKSGRWVMLKNVHLAP GWLMQLEKKLHSLQPHACFRLFLTMEINPKVPV NLLRAGRIFVFEPPPGVKANMLRTFSSIPVSRICK SPNERARLYFLLAWFHAIIQERLRYAPLGWSKKY EFGESDLRSACDTVDTWLDDTAKGRQNISPDKIP WSALKTLMAQSIYGGRVDNEFDQRLLNTFLERL FTTRSFDSEFKLACKVDGHKDIQMPDGIRREEFV QWVELLPDTQTPSWLGLPNNAERVLLTTQGVD MISKMLKMQMLEDEDDLAYAETEKKTRTDSTS DGRP\AWMRTLHTTASNWLHLIPQTLSHLKRTVE NIKDPLFRFFE\REVKMGAKLLQ\DVRQDLADV\V QVCEGKKKQTNYLRTLI\NELV\KGILP\RSWSHY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TVPAG\MTVIQWGVPISARRI\KQLQNISL\AAASGGAKELKNIHVCLGGLFVPEAYITATRQYVAQANSWSLEELCLEVNVTTSQGATLDACSFGVTGLKLQGATCNNNKLSLSNAISTALPLTQLRWVKQTNTEKKASVVTLPVYLNFTRADLIFTVDFEIATKEDPRSFYERGVAVLCTE
3771	A	1	2043	LPLLHAGFNRRFMENSSIIACYNELIQIEHGEVRSQFKLRACNSVFTALDHCHEAIEITSDDHVIQYVNPAFERMMGYHKGELLGKELADLPKSDKNRADLLDTINTCIKKGKEWQGVYYARRKSGDSIQQHVKITPVIGQGGKIRHFVSLKKLCCTTDNNKQIHKIHRDSGDNSQTEPHSFRYKNRRKESIDVKSISSRGSDAPSLQNRRYPSMARIHSMTIEAPITKVINIINAAQENSPVTVAEALDRVLEILRTTELYSPQLGTKDEDPHTSDLVGGLMTDGLRRLSGNEYVFTKNVHQSHSHLAMPITINDVPPCISQLLDNEESWDFNIFELEAITHKRPLVYLGLKVFSRFGVCEFLNCSETTLRAWFQVIEANYHSSNAYHNSTHAADVLHATAFFLGKERVKGSLDQLDEVAALIAATVHDVDHPGRTNSFL\CNAGSELAVLYNDT\AV\LESHHTALAFQ\LTVKDTK\CNIFKNID/RGNHYRTLRQAIIDMVLATEMTKHFEHVNKFVNSINKPMAAEIEGSDCECNPAGKNFPENQILIKRMMIKCADVANPCRPLDLCIEWAGRISEEYFAQTDEEKRQGLPVVMPVFDRNTCSIPKSQISFIDYFITDMFDAWDAFAHLPALMQHLADNYKHWKTLDDLKCKSLRLPSDRLKPSHRGGLLTDKGHCESQ
3772	A	1013	50	TLVHADGFPSLHITETCLAYREKRIGIDLVHDTVEHELIKEAEIIQGIMALLTRTLEEASEQIRMNRSAKYNLEKDLKDKFVALTIDDICFSLNNNSPNIRYSENAVRIEPNSVSLEDWLDFSSTNVEKADKQRNNSLMLKALVD\RILSQTANYLRKQCDVVHTAFKNGLKDTKDARDQLADHLAK\VMEEIASQEKNITALEKAILDQEGPAKVAHTRLETRTHRPNVELCRDVAQYRLMKEVQEITHNVARLKETLA\QAQAELKGLHRRQLALQEEIQVKENTIYIDEVLCMQMRKSIPLRDGEDHGVWAGGLRPDAVC
3773	A	1	955	AAARESERQLRLRLCVLNEILGTERDYVGTLRFLQSAFLHRIRQNVADSVEKGLTEENVKVLFSNIEDILEVHKDFLAALEYCLHPEPQSQHELGNVFLKFKDKFCVYEEYCSNHEKALRLLVELNKIPTVRAFLLSCMLLGGRKTTDIPLEGYL\LSPIQRICKYPLLLKELAKRTPGKHPDHPAVQ\SALQAMKTVCSNINETKRQMEKLEALEAAA/QSHIEGWEGSNLTDICTQLLLQGTLLKISAGNIQERAFFLFDNLLVYCKRKSRVTGSKKSTKRTKSINGSLYIFRGRINTEVMEVENVEDGTGSPSPSLA
3774	A	4254	2061	ELQGDFSVPDVPKSMAWCENSICVGFKRDYYLIRVDGKGSIKELFPTGKQLEPLVAPLADGKVAVGQDDLTVVLNEEGICTQKCALNWTDIPVAMEHQPPYIIAVLPRYVEIRTFEPRLLVQSIELQRPRFITSGGSNIIYVASNHFVWRLIPVPMATQIQQLLQDKQFELALQLAEMKDDSDSEKQQQIHHIKNLYAFNLFCQKRFDESMQVFAKLGTDPTHVMGLYPDLLPTDYRKQLQYPNPLPVLSGAELEKAHLALIDYLTQKRS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QLVKKLNDSDHQSSTSPLMEGTPTIKSKKKLLQII DTTLLKCYLHTNVALVAPLLRLENNHCHIEESEH VLKKAHKYSELIILYEKKGLHEKALQVLVDQSK KANSPLKGHERTVQYLQHLGTENLHLIFSYSVW VLRDFPEDGLKIFTEDLPEVESLPRDRVLGFLIEN FKGLAIPYLEHIIHVWEETGSRFHNCLIQLYCEKV QGLMKEYLLSFPAGKTPVPAGEEEGELGEYRQK LLMFLEISSYYDPGRLICDFPFDGLLEERALLLGR MGKHEQALFIYVHILKDTRMAEEYCHKHYDRN KDGNKDVYLSLLRMYLSPPSIHCLGPIKLELLEPK ANLQAALQVLELHHSKLDTTKALNLLPANTQIN DIRIFLEKVLEENAQKKRFNQVLKNLLHAEFLRV\ QEERILHQQVKCIITEEKVCMVCKKKIGNSAFAR YPNGVVVHYFCS\KEVNPADT
3775	A	1832	839	MSRARGALCRACLALAAALAALLLLPLPLPRAP APARTPAPAPRAPPSRPAAPSLRPDDVFIAVKTTR KNHGPRLRLLLRTW\ISRARQQTFIFTDGDDPELE LQGGDRVINTNCSAVRTRQALCCKMSVEYDKFI ESGRKWFCHVDDDNYVNARSLLHLLSSFSPSQD VYLGRPSLDHPIEATERVQGGRTVTTVKFWFAT GGAGFCLSRGLALKMSPWASLGSFMSTAEQVRL PDDCTVGYIVEGLLGARLLHSPLFHSHLENLQRL PPDTLLQQVTLSHGGPENPQNVVNVAGGFSLHQ DPTRFKSIHCLLYPDTDWCPRQKQGAPTSR
3776	A	3	796	PRAKLGTRARNMAGQDAGCGRGGDDYSEDEGD SSVSRAAVEVFGKLKDLNCPFLEGLYITEPKTIQE LLCSPSEYRLEILEWMCTRVWPSLQDRFSSLKGV PTEVKIQEMTKLGHELMLCAPDDQELLKGCACA QKQLHFMDQLLDTIRSLTIGCSSCSSLMEHFEDT REKNEALLGELFSSPHLQMLLNPECDPWPLDMQ PLLNKQSDDWQWASASAKSEEEEKLAELARQLQ ESAAKLHALRTEYFAQHEQGAAAGAA\TSAP
3777	A	3	413	SEEDVIEGKTAVIEKRRKKRSSAGVVED/IGGEVQ NMLEGVGVDINKALLAKRKRLEMYTKASLRTSN QKIEHVWKTQQDQRQKLNQEYSQQFLTLFQQW DLDMQKAEEQEEKILVGIMIRFIINQVSSRNGQPS LLL
3778	A	132	788	SRLPPPPPHLADGRAGARVPRSARLSRWWVQD WTHGPIVRPPAAARTMWVNPEEVLLANALWITE RANPYFILQRRKGHAGDGGGGGGLAGLLVGTLD VVLDSSARVAPYRILYQTPDSLVYWTIACG\GSR KEITEHWEWLEQNLLQTLSIFENENDITTFVRGKI QGIIAEYNKINDVKEDDDTEKFKEAIVKFHRLFG MPEEEKLVNYYSCSYWKG
3779	A	2	934	CKSCTLFPQNPNLPPPSTRERPPGCKTVFVGGLPE NATEEIIQEVFEQCGDITAIRKSKKNFCHIRFAEEF MVDKAIYLSGYRMRLGSSTDKKDSGRLHVDFA QARDDFYEWECKQRMRAREERHRRKLEEDRLR PPSPPAIMHYSEHEAALLAEKLKDDSKFSEAM\Q VLLSWIERGEVNRR\SANQFYSMVQSANSHVRRL MNEKATHEQEMEEAKENFKNALTGILTQFEQIV AVFNASTRQKAWDHFSKAQRKNIDIWAK\HSEE LRNAQSEQLMGIRREEEMEMSDDENCDSPTKKM RVDESALGAP
3780	A	1	2535	AAQAEREELAAGRMPGGGPQGAPAAAGGGGVS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HRAGSRDCLPPAACFRRRRLARRPGYMRSSTGPGIGFLSPAVGTLFRFPGGVSGEESHHSESRARQCGLDSRGLLVRSPVSKSAAAPTVTSVRGTSAHFGIQLRGGTRLPDRLSWPCGPGSAGWQQEFAAMDSSETLDASWEAACSDGARRVRAAGSLPSAELSSNSCSPGCGPEVPPTPPGSHSAFTSSFSFIRLSLGSAGERGEAEGCPPSREAESHCQSPQEMGAKAASLDGPHEDPRCLSQPFSLLATRVSADLAQAARNSSRPERDMHSLPDMDPGSSSSLDPSLAGCGGDGSSGSGDAHSWDTLLRKWEPVLRDCLLRNRRQMEVISLRLKLQKLQEDAVENDDYDKAETLQQRLEDLEQEKISLHFQLPSRQPALSSFLGHLAAQVQAALRRGATQQASGDDTHTPLRMEPRLLEPTAQDSLHVSITRRDWLLQEKQQLQKEIEALQARMFVLEAKDQQLRREIEEQEQQLQWQGCDLTPLVGQLSLGQLQEVSKALQDTLASAGQIPFHAEPPETIRSLQERIKSLNLSLKEITTKVCMSEKFCSTLRKKVNDIETQLPALLEAKMHAISGNHFWTAKDLTEEIRSLTSDREGLEGLLSKLLVLSSRNVKKLGSVKEDYNRLRREVEHQETAYETSVKENTMKYMETLKNKLCSCKCPLLGKVWEADLEACRLLIQCLQLQEARGSLSVEDERQMDDLEGAAPPIPPRLHSEDKRKTPLKESYILSAELGEKCEDIGKKLLYLEDQLHTAIHSHDEDLIQSLRRELQMVKETLQAMILQLQPAKEAGEREAAASCMTAGVHEAQA
3781	A	3	995	GRRRAGPAHSARMYNMMETELKPPGPQQTSGGGGGNSTAAAAGGNQKNSPDRVKRPMNAFMVWSRGQRRKMAQENPKMHNSEISKRLGAEWKLLSETEKRPFIDEAKRLRALHMKEHPDYKYRPRRKTKTLMKKDKYTLPGGLLAPGGNSMASGVGVGAGLGAGVNQRMDSYAHMNGWSNGSYSMMQDQLGYPQHPGLNAHGAAQMQPMHRYDVSALQYNSMTSSQTYMNG/SRPTYSMSYSQQGTPGMAPGS\MGSVVKSEASSSPPVVTSSSHSRAPCQAGDLRDMISMYLPGAEVPEPAAPSRLHMSQHYQSGPVPGTAINGTLPLSHM
3782	A	1	2649	FRVPDSCPVVLHSFTQLDPDLPRPESSTQEIGEELINGVIYSISLRKVQLHHGGNKGQRWLGYENESALNLYETCKVRTVKAGTLEKLVEHLVPAFQGSDLSYVTIFLCTYRAFTTTQQVLDLLFKRYGRCDALTASSRYGCILPYSDEDGGPQDQLKNAISSILGTWLDQYSEDFCQPPDFPCLKQLVAYVQLNMPGSDLERRAHLLLAQLEHSEPIEAEPEGEEDWALSPVPALKPTPELELALTPARAPSPVPAPAPEPEPAPTPAPGSELEVAPAPAPELQQAPEPAVGLESAPAPALELEPAPEQDPAPSQTLELEPAPAPVPSLQPSWPSPVVAENGLSEEKPHLLVFPPDLVAEQFTLMDAELFKKVVPYHCLGSIWSQRDKKGKEHLAPTIRATVTQFNSVANCVITTCLGNRSTKAPDRARVVEHWIEVARECRILKNFSSLYAILSALQSNSIHRLKKTWEDVSRDSFRIFQKLSEIFSDENNYSLSRELLIKEGTSKFATLEMNPKRAQKRPKETGIIQGTVPYLGTFLTDLVMLDTAMKDYLYGRLINFEKRRKEFEVIAQIKLLQSACNNYSIAPDEQFGAWFRAVERLSETESYNLSCELEPPSESASNTLRTKKNTAIVKRWSDRQAPSTELS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				TSGSSHSKSCDQLRCGPYLSSGDIADALSVHSAG SSSSDVEEINISFVPESPDGQEKKFWESASQSSPET SGISSASSSTSSSSASTTPVAATRTHKRSVSGLCNS SSALPLYNQQVGDCCIIRVSLDVDNGNMYKSILV TSQDKAPAVIRKAMDKHNLEEEEPEDYELLQILS DDRKLKIPENANVFYAMNSTANYDFVLKKRTFT KGVKVKHGASSTLPRMKQKGLKIAKGIF
3783	A	3	869	RSGQGKVYGLIGRRRFQQMDVLEGLNLLITISGK RNKLRVYYLSWLRNKILHNDPEVEKKQGWTTV GDMEGCGHYRVVKYERIKFLVIALKSSVEVYAW APKPYHKFMAFKSFADLPHRPLLVDLTVEEGQR LKVTYGSSAGFHAVDVDSGNSYDIYIPVHIQSQIT PHAIIFLPNTDGMEMLLCYEDEGVYVNTYGRIIK DVVLQWGEMPTSVAYICSNQIMGWGEKAIEIRS VETGHLDGVFMHKRAQRLKFLCERNDKVFFASV RSGGSSQVYFMTLNRNCIMNW
3784	A	1213	457	LSPRQVDGLAGLQKGLSLSLLYQFLMNGIRLGTY GLAEAGGYLHTAEGTHSPARSAAAGAMAGVMG AYLGSPIYMVKTHLQAQAASEIAVGHQYKHQG MFQALTEIGQKHGLVGLWRGALGGLPRVIVGSS TQLCTFSSTKDLLSQWEIFPPQSWKLALVAAMM SGIAVVLAMAPFDVACTRLYNQPHRCTGQGP\LY RGILDALLQTARTEGIFGMYKGIGASYFRLGPHTI LSLFFWDQLRSLYYTDTK
3785	A	193	813	RRRGRHSLCGGKMLAYCVQDATVVDVEKRRNP SKHYVYIINVTWSDSTSQTIYRRY\SKFFDLQMQL LD\KFPI\ESGQKDPKQRIIPFLPGKILFRRSHIRDV AVKRLKPIDEYCRALVRLPPHISQCDEVFRFFEAR PEDVNPPKEQGPSPPDAVLPYGVNKGKQELKAG PNWPGRTHHVVNCVTQKCLFVFHFKFSSSGNKE SKSL
3786	A	3785	1632	EFVGRAASTTVVTRIAWRMADAGIRRVVPSDLY PLVLGFLRDNQLSEVANKFAKATGATQQDANAS SLLDIYSFWLNRSAKVPERKLQANGPVAKKAKK KASSSDSEDSSEEEEEVQGPPAKKAAVPAKRVGL PPGKAAAKASESSSSEESSDDDDEEDQKKQPVQ KGVKPQAKAGQAPPKKAKSSDSDSDSSSEDEPP KNQKPKITP\VTVKAQTKAPPKPARA\APKIANGK AASSSSSSSSSSSSDDSEEEKAAATPKKTVPKKQV VAKAPVKAATTPTRKSSSSEDSSSDEEEEQKKPM KNKPGPYSSVPPPSAPPPKKSLGTQPPKKAVEKQ QPVESSEDSSDESDSSSEEEKKPPTKAVVSKATTK PPPAKKAAESSSDSSDSDSSEDDEAPSKPAGTTK NSSNKPAVTTKSPAVKPAAAPKQPVGGGQKLLT RKADSSSSEEESSSSEEEKTKKMVATTKPKATAK AALSLPAKQAPQGSRDSSSDSDSSSSEEEEEKTSK SAVKKKPQKVAGGAAPSKPASAKKGKAESSNSS SSDDSSEEEEEKLKGKGSPRPQAPKANGTSALTA QNGKAAKNSEEEEEEKKKAAVVVSKSGSLKKR KQNEAAKEAETPQAKKIKLQTPNTFPKRKKGEK RASSPFRRVREEEIEVDSRVADNSFDAKRGAAGD WGERANQVLKFTKGKSFRHEKTKKKRGSYRGG SISVQVNSIKFDSE
3787	A	3	5078	IPEG/RALSAEHTSSLVPSLHITTLGQEQAILSGAV PASPSTGTADFPSILTFLQPTENHASPSPVPEMPTL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				PAEGSDGSPPATRDLLLSSKVPNLLSTSWTFPRW KKDSVTAILGKNEEANVTIPLQAFPRKEVLSLHT VNGFVSDFSTGSVSSPIITAPRTNPLPSGPPLPSILS IQATQTVFPSLLAFSSTKPEVYAAAVDHSGLPAS APKQVRASPSSMDVYDSLTIGDMKKPATTDVFW SSLSAETGSLSTESIISGLQQQTNYDLNGHTISTTS WETHLAPTAPPNGLTSAADAIKSQDFKDTAGHS VTAEGFSIQDLVLGTSIEQPVQQSDMTMVGSHID LWPTSNNNHSRDFQTAEVAYYSPTTRHSVSHPQ LQLPNQPAHPLLLTSPGPTSTGSLQEMLSDGTDT GSEISSDINSSPERNASTPFQNILGYHSAAESSISTS VFPRTSSRVLRASQHPKKWTADTVSSKVQPTAA AAVTLFLRKSSPPALSAALVAKGTSSSPLAVASG PAKSSSMTTLAKNVTNKAASGPKRTPGAVHTAF PFTPTYMYARTGHTTSTHTA/IARKHGHCLWPVV YNLP/PP/GKPQAMHTGLPNPTNLEMPRASTPRPL TVTAALTSITASVKATRLPPLRAENTDAVLPAAS AAVVTTGKMASNLECQMSSKLLVKTVLFLTQRR VQISESLKFSIAKGLTQALRKAFHQNDVSAHVDI LEYSHNVTVGYYATKGKLVYLPAVVIEMLGVY GVSNVTADLKQHTPHLQSVAVLASPWNPQPAG YFQLKTVLQFVSQADNIQSCKFAQTMEQRLQKA FQDAERKVLNTKSNLTIQIVSTSNASQAVTLVYV VGNQSTFLNGTVASSLLSQLSAELVGFYLTYPPL TIAEPLEYPNLDISETTRDYWVITVLQGVDNSLV GLHNQSFARVMEQRLAQLFMMSQQQGRRFKRA TTLGSYTVQMVKMQRVPGPKDPAELTYYTLYN GKPLLGTAAAKILSTIDSQRMALTLHHVVLLQAD PVVKNPPNNLWIIAAVLAPIAVVTVIIIIITAVLCR KNKNDFKPDTMINLPQRAKPVQGFDYAKQHLG QQGADEEVIPVTQETVVLPLPIRDAPQERDVAQD GSTIKTAKSTETRKSRSPSENGSVISNESGKPSSGR RSPQNVMAQQKVTKEEARKRNVPASDEEEGAV LFDNSSKVAAEPFDTSSGSVQLIAIKPTALPMVPP TSDRSQESSAVLNGEVNKALKQKSDIEHYRNKL RLKAKRKGYYDFPAVETSKGLTERKKMYEKAP KEMEHVLDPDSELCAPFTESKNRQQMKNSVYRS RQSLNSPSPGETEMDLLVTRERPRRGIRNSGYDT EPEIIEETNIDRVPEPRGYSRSRQVKGHSETSTLSS QPSIDEVRQQMHMLLEEAFSLASAGHAGQSRHQ EAYGSAQHLPYSEVVTSAPGTMTRPRAGVQWVP TYRPEMYQYSLPRPAYRFSQLPEMVMGSPPPPVP PRTGPVAVASLRRSTSDIGSKTRMAESTGPEPAQ LHDSASFTQMSRGPVSVTQLDQSALNYSGNTVP AVFAIPAANRPGFTGYFIPTPPSSYRNQAWMSYA GENELPSQWADSVPLPGYIEAYPRSRYPQSSPSRL PRQYSQPANLHPSLEQAPAPSTAASQQSLAENDP SDAPLTNISTAALVKAIREEVAKLAKKQTDMFEF QV
3788	A	2	1737	MKGLYTDAEMKSDNVKDKDAKISFLQKAIDVV VMVSGEPLLAKPARIVAGHEPERTNELLQIIGKC CLNKLSSDDAVRRVLAGEKGEVKGRASLTSRSQ ELDNKNVREEESRVHKNTEDRGDAEIKERSTSRD RKQKEELKEDRMPREKDKDKEKAKENGGNRHR EGERERAKARARPDNERQKDRGNRERDRDSERK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KETERKSEGGKEKERLRDRDRERDRDKGKDRDR RRVKNGEHSWDLDRENNREHDKPEKKSASSGE MSKKLSDGTFKDSKAETETEISTRASKSLTTKTS KRRSKNSVEGDSTSDAEGDAGPAGQDKSEVPET PEIPNELSSNIRRIPRPGSARPAPPRVKRQDSMEAL QMDRSGSGKTVSNVITESHNSDNEEDDQFVVEA APQLSEMSEIEMVTAVELEEEEKHGGLVKKILET KKDYEKLQQSPKPGEKERSLFESAWKKEKDIVS KEIEKLRTSIQTLCKSALPLGKIMDYIQEDVDAM QNELQM\YHSENRQHAEALQQEQRITDCAVEP\L KAELA\ELEQLIKD\Q\QDKICAVKANILKNEEKIQ KMVYSINLTSRR
3789	A	1	4369	MRTLGTCLATLAGLLLTAAGETFSGGCLFDEPYS TCGYSQSEGDDFNWEQVNTLTKPTSDPWMPSGS FMLVNASGRPEGQRAHLLLPQLKENDTHCIDFH YFVSSKSNSPPGLLNVYVKVNNGPLGNPIWNISG DPTRTWNRAELAISTFWPNFYQVIFEVITSGHQG YLAIDEVKVLGHPCTRTPHFLRIQNVEVNAGQFA TFQCSAIGRTVAGDRLWLQGIDVRDAPLKEIKVT SSRRFIASFNVVNTTKRDAGKYRCMI\RTEGGVGI SNYAEL\VVKEPPVPIAPPQLASVGATYLWIQLN ANSINGDGPIVAREVEYCTASGSWNDRQPVDSTS YKIGHLDPDTEYEISVLLTRPGEGGTGSPGPALRT RTKCADPMRGPRKLEVVEVKSRQITIRWEPFGY NVTRCHSYNLTVHYCYQVGGQEQVREEVSWDT ENSHPQHTITNLSPYTNVSVKLILMNPEGRKESQ ELIVQTDEDLPGAVPTESIQGSTFEEKIFLQWREP TQTYGVITLYEITYKAVSSFDPEIDLSNQSGRVSK LGNETHFLFFGLYPGTTYSFTIRASTAKGFGPPAT NQFTTKISAPSMPAYELETPLNQTDNTVTVMLKP AHSRGAPVSVYQIVVEEERPRRTKKTTEILKCYP VPIHFQNASLLNSQYYFAAEFPADSLQAAQPFTIG DNKTYNGYWNTPLLPYKSYRIYFQAASRANGET KIDCVQVATKGAATPKPVPEPEKQTDHTVKIAG VIAGILLFVIIFLGVVLVMKKRKL\AKKRKETMSS TRQEIDLWIGELNGPRSYAEQGTKLATRAFSFMD THNLNGRSVSSPSSFTMKTNTLSTSVPNSYYPDE THTMASDTSSLVQSHTYKKREPADVPYQTGQLH PAIRVADLLQHITQMKCAEGYGFKEEYESFFEGQ SAPWDSAKKDENRMKNRYGNIIAYDHSRVRLQT IEGDTNSDYINGNYIDGYHRPNHYIATQGPMQET IYDFWRMVWHENTASIIMVTNLVEVGRVKCCK YWPDDTEIYKDIKVTLIETELLAEYVIRTFAVEKR GVHEIREIRQFHFTGWPDHGVPYHATGLLGFVR QVKSKSPPSAGPLVVHCSAGAGRTGCFIVIDIML DMAEREGVVDIYNCVRELRSRRVNMVQTEEQY VFIHDAILEACLCGDTSVPASQVRSLYYDMNKLD PQTNSSQIKEEFRTLNMVTPTLRVEDCSIALLPRN HEKNRCMDILPPDRCLPFLITIDGESSNYINAALM DSYKQPSAFIVTQHPLPNTVKDFWRLVLDYHCTS VVMLNDVDPAQLCPQYWPENGVHRHGPIQVEF VSADLEEDIISRIFRIYNAARPQDGYRMVQQFQFL GWPMYRDTPVSKRSFLKLIRQVDKWQEEYNGG EGRTVVHCLNGGGRSGTFCAISIVCEMLRHQRTV DVFHAVKTLRNNKPNMVDLLDQYKFCYEVALE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				YLNSG
3790	A	261	485	EEQTPLHIASRLGKTEIVQLLLQHMAHPDAATTN GYTPLHISAREGQV\DV\ASVLLGRQGAAHSFRLT KVRRMTS
3791	A	1	5874	LPPVTMSGKYIMEEHDSYSDQVWSIDELPSKQG YYLQGNYLRCVAEVGSFEHNLTTDLLNHLVFVQ KVFMKEVNEVIQKVSGGEQPIPLWNEHDGTADG DKPKILLYSLNLQFKGIQVTATTPSMRAVRFETG LIELELSNRLQTKASPGSSSYLKLFGKCQVDLNL ALGQIVKHQVYEEAGSDFHQVAYFKTRIGLRNA LREEISGSSDREAVLITLNRPIVYAQPVAFDRAVL FWLNYK\AAYDNWNEQRMALHKDIHMATKEVV DMLPGIQQTSAQAFGTPFLQLTVNDLGICLPITNT AQSNHTGDLDTGSALVLTIESTLITACSSESLVSK GHFKNFCIRFADGFETSWDDWKPEIHGDLVMNA CVVPDGTYEVCSRTTGQAAAESSSAGTWTLNVL WKMCGIDVHMDPNIGKRLNALGNTLTTLTGEED IDDIADLNSVNIADLSDEDEVDTMSPTIHTEATDY RRQAASASQPGELRGRKIMKRIVDIRELNEQAKV IDDLKKLGASEGTINQEIQRYQQLESVAVNDIRR DVRKKLRRSSMRAASLKDKWGLSYKPSYSRSKS ISASGRPPLKRMERASSRVGETEELPEIRVDAASP GPRVTFNIQDTFPEETELDLLSVTIEGPSHYSSNSE GSCSVFSSPKTPGGFSPGIPFQTEEGRRDDSLSSTS EDSEKDEKDEDHERERFYIYRKPSHTSRKKATGF AAVHQLFTERWPTTPVNRSLSGTATERNIDFELD IRVEIDSGKCVLHPTTLLQEHDDISLRRSYDRSSR SLDQDSPSKKKKFQTNYASTTHLMTGKKVPSSL QTKPSDLETTVFYIPGVDVKLHYNSKTLKTESPN ASRGSSLPRTLSKESKLYGMKDSATSPPSPPLPST VQSKTNTLLPPQPPPIPAAKGKGSGGVKTAKLYA WVALQSLPEEMVISPCLLDFLEKALETIPITPVER NYTAVSSQDEDMGHFEIPDPMEES\TTSLVS\SSTS AYSSFPVDVVVYVRVQPSQIKFSCLPVSRVECML KLPSLDLVFSSNRGELETLGTTYPAETLSPGGNA TQSGTKTSASKTGIPGSSGLGSPLGRSRHSSSQSD LTSSSSSSSGLSFTACMSDFSLYVFHPYGAGKQIT AVSGLTPGSGGLGNVDEEPTSVTGRKDSLSINLE FVKVSLSRIRRSGGASFFESQSVSKSASKMDTTLI NISAVCDIGSASFKYDMRRLSEILAFPRAWYRRSI ARRLFLGDQTINLPTSGPGTPDSIEGVSQHLSPESS RKAYCKTWEQPSQSASFTHMPQSPNVFNEHMTN STMSPGTVGQSLKSPASIRSRSVSDSSVPRRDSLS KTSTPFNKSNKAASQQGTPWETLVVFAINLKQL NVQMNMSNVMGNTTWTTSGLKSQGRLSVGSNR DREISMSVGLGRSQLDSKGGVVGGTIDVNALEM VAHISEHPNQQPSHKIQITMGSTEARVDYMGSSIL MGIFSNADLKLQDEWKVNLYNTLDSSITDKSEIF VHGDLKWDIFQVMISRSTTPDLIKIGMKLQEFFT QQFDTSKRALSTWGPVPYLPPKTMTSNLEKSSQE QLLDAAHHRHWPGVLKVVSGCHISLFQIPLPEDG MQFGGSMSLHGNHMTLACFHGPNFRSKSWALF HLEEPNIAFWTEAQKIWEDGSSDHSTYIVQTLDF HLGHNTMVTKPCGALESPMATITKITRRRHENPP HGVASVKEWFNYVTATRNEELNLLRNVDANNT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ENSTTVKNSSLLSGFRGGSSYNHETETIFALPRMQLDFKSIHVQEPQEPSLQDASLKPKVECSVVTEFTDHICVTMDAELIMFLHDLVSAYLKEKEKAIFPPRILSTRPGQKSPIIIHDDNSSDKDREDSITYTTVDWRDFMCNTWHLEPTLRLISWTGRKIDPVGVDYILQKLGFHHARTTIPKWLQRGVMDPLDKVLSVLIKKLGTALQDEKEKKGKDKEEH
3792	A	1	364	QNGSTPLHHAASKNRHEIALMLLEGGANPDGKDHYEATAKHQATAKGNFKMIHILLYYKASTIIQDTEGNTPPHLVCD\RVEEAKLLVSQGA/SIYIENKEEKDP/LQVAKGALGLVLKRMVEG
3793	A	2	340	DIVPNPKMAPLGDEAPTLEKVLTPELSEEEVSTRDDIQFHHFSSEEALQKVKYFVAKEDPSSQEEAHTPEAPPPQPPSSERCLGEMKCTLVRGDSSPRQAELKSGPASRPAL
3794	A	421	158	SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQWISKAVYKHREMCGLTSTGRKSHGLEKDRMFPHAIGGSCRAA*RRRKTLQFPCYH
3795	A	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDGTLERGHWNNKMEFVLSVAGEIIGLGNVWRFPYLCYKNGGGAFFIPYLVFLFTCGIPVFLLETALGQYTSQGGVTAWRKICPIFEGIGYASQMIVILLNVYYIIVLAWALFYLFSSFTIDLPWGGCYHEWNTEHCMEFQKTNGSLNGTSENATSPVIEFW
3796	A	3	592	KPASTYSTSQPSMAPLLPIRTLPLILILLALLSPGAADFNISSLSGLLSPALTESLLVALPPCHLTGGNATLMVRRANDSKVVTSSFVVPPCRGRRELVSVVDSGAGFTVTRLSAYQVTNLVPGTKFYISYLVKKGTATESSREIPMFTLPRRNMESIGLGMARTGGMVVITVLLSVAMFLLVLGFIIALALGSRK
3797	A	1	1556	ATRLLRGSGSWGCSRLRFGPPAYRRFSSGGAYPNIPLSSPLPGVPKPVFATVDGQEKFETKVTTLDNGLRVASQNKFGQFCTVGILINSGSRYEAKYLSGIAHFLEKLAFSSTARFDSKDEILLTLEKHGGICDCQTSRDTTMYAVSADSKGLDTVVALLADVVLQPRLTDEEVEMTRMAVQFELEDLNLRPDPEPLLTEMIHEAAYRENTVGLHRFCPTENVAKINREVLHSYLRNYYTPDRMVLAGVGVEHEHLVDCARKYLLGVQPAWGSAEAVDIDRSVAQYTGGIAKLERDMSNVSLGPTPIPELTHIMVGLESCSFLEEDFIPFAVLNMMMGGGGSFSAGGPGKGMFSRLYLNVLNRHHWMYNATSYHHSYEDTGLLCIHASADPRQVREMVEIITKEFILMGGTVDTVELERAKTQLTSMLMMNLESRPVIFEDVGRQVLATRSRKLPHELCTLIRNVKPEDVKRVASKMLRGKPAVAALGDLTDLPTYEHIQTALSSKDGRLPRTYRLFR
3798	A	73	759	KRLVEAGVPRTFDGIVGEGGAQSRSCWPWGVTAQTPAFSADSLNCLKNCMSITMGSVRPSVEQFHKYLPWFLNDRPNIKCPKGGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYTITNVFYEQYLTILPEGLFMLSLCLVPTFAVSCLLLGLDLRSGLLNLLSIVMILVDTVGFMALWGISYNAVSLINLVS
3799	A	73	759	KRLVEAGVPRTFDGIVGEGGAQSRSCWPWGVTAQTPAFSADSLNCLKNCMSITMGSVRPSVEQFHKY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LPWFLNDRPNIKCPKGGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYTITNVFYEQYLTILPEGLFMLSLCLVPTFAVSCLLLGLDLRSGLLNLLSIVMILVDTVGFMALWGISYNAVSLINLVS
3800	A	250	1032	GIFRSLRVLFPLFSVGRPQFARSLSAAPQLSDTADTMGFGDLKSPAGLQVLNDYLADKSYIEGYVPSQADVAVFEAVSSPPPADLCHALRWYNHIKSYEKEKASLPGVKKALGKYGPADVEDTTGSGATDSKDDDDIDLFGSDDEEESEEAKRLREERLAQYESKKAKKPALVAKSSILLDVKPWDDETDMAKLEECVRSIQADGLVWGSSKLVPVGYGIKKLQIQCVVEDDKVGTDMLEEQITAFEDYVQSMDVAAFNKI
3801	A	155	656	SREMELVTFRDVAIEFSPEEWKCLDPAQQNLYRDVMLENYRNLVSLGFVISNPDLVTCLEQIKEPCNLKIHETAAKPPAICSPFSQDLSPVQGIEDSFHKLILKRYEKCGHENLQLRKGCKRVNECKVQKGVNNGVYQCLSTTQSKIFQCNTCVRVFSTSSHSNKHK
3802	A	1	1428	VTVSPETHMDLTKGCVTFEDIAIYFSQDEWGLLDEAQRLLYLEVMLENFALVASLGCGHGTEDEETPSDQNVSVGVSQSKAGSSTQKTQSCEMCVPVLKDILHLADLPGQKPYLVGECTNHHQHQKHHSAKKSLKRDMDRASYVKCCLFCMSLKPFRKWEVGKDLPAMLRLLRSLVFPGGKKPGTITECGEDIRSQKSHYKSGECGKASRHKHTPVYHPRVYTGKKLYECSKCGKAFRGKYSLVQHQRVHTGERPWECNECGKFFSQTSHLNDHRRIHTGERPYECSECGKLFRQNSSLVDHQKIHTGARPYECSQCGKSFSQKATLVKHQRVHTGERPYKCGECGNSFSQSAILNQHRRIHTGAKPYECGQCGKSFSQKATLIKHQRVHTGERPYKCGDCGKSFSQSSILIQHRRIHTGARPYECGQCGKSFSQKSGLIQHQVVHTGERPYECNKCGNSFSQCSSLIHHQKCHNT
3803	A	193	617	LFPFLGSESKNGEADSSDKEMKHGQKSPTGKQTSQHLKRLKKSGLGHLKWTKAEDIDIETPGSILVNTNLRALINKHTFASLPQHFQQYLLLLLPEVDRQMGSDGILRLSTSALNNEFFAYAAQGWKQRLAEGKFVFSIIM
3804	A	197	479	SSSRASPPEHPSSQAHCGPLVLSHACPEVTNKWSTGSSSSPNSSWVSSPLQPEGLSGSSRMKGGSATKILLETLLLAAHMTADQGIASSQRCLL
3805	A	1	385	QSADTLFPGDINFNVSGLFSAVTLQDTVSDRLASEELPSTAVPTPATTPAPAPAPAPATAPALVSAATKERTESEVPPRPASPKVTRSPPETAAPVEDMARRSELAVGGEEGTEGGRGEGTGSPMSSY
3806	A	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNLLQPQAPGHDMTSIPFPGDRLLQVDGVILCGLTHKQAVQCLKGPGQVARLVLERRVPRSTQQCPSANDSMGDERTAVSLVTALPGRPSSCVSVTDGPKF*SSN*KRIANGLGFSFVQMEKESCSHLKSDLVRIKRLFPGHPAEENGAIAAGDIILGREWEGPRKASSSRCRGSWAMQLSVQAGPSFASYYPAAVEVLHLLRGAPQEVTLLLCRPPPGALPELEQEWQTPELSADKEFTRATCTDSCTSPILGSRGQLGGTVPPQMQGKAWGLRPESSQKAIREGTMGAKTERDLGPVP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3807	A	656	1238	RCPSLLPPSWPLPTLQTLTRTPGNKAIAGGAGLWAVLWGSERTPPYR*GN*NQRGAVPCLRPHRLRPQDKFLVLASDGLWDMLSNEDVVRLVVGHLAEADWHKTDLAQRPANLGLMQSLLLQRKASGLHEADQNAATRLIRHAIGNNEYGEMEAERLAAMLTLPEDLARMYRDDITVTVVYFNSESIGAYYKGG
3808	A	26	2195	SQYSESVAGRQASPERLLGSYHAMASTVEGGDTALLPEFPRGPLDAYRARASFSWKELALFTEGEGMLRFKKTIFSALENDPLFARSPGADLSLEKYRELNFLRCKRIFEYDFLSVEDMFKSPLKVPALIQCLGMYDSSLAAKYLLHSLVFGSAVYSSGSERHLTYIQKIFRMEIFGCFALTELSHGSNTKAIRTTAHYDPATEEFIIHSPDFEAAKFWVGNMGKTATHAVVFAKLCVPGDQCHGLHPFIVQIRDPKTLLPMPGVMVGDIGKKLGQNGLDNGFAMFHKVRVPRQSLLNRMGDVTPEGTYVSPFKDVRQRFGASLGSLSSGRVSIVSLAILNLKLAVAIALRFSATRRQFGPTEEEEIPVLEYPMQQWRLLPYLAAVYALDHFSKSLFLDLVELQRGLASGDRSARQAELGREIHALASASKPLASWTTQQGIQECREACGGHGYLAMNRLGVLRDDNDPNCTYEGDNNILLQQTSNYLLGLLAHQVHDGACFRSPLKSVDFLDAYPGILDQKFEVSSVADCLDSAVALAAYKWLVCYLLRETYQKLNQEKRSGSSDFEARNKCQVSHGRPLALAFVELTVVQRFHEHVHQPSVPPSLRAVLGRLSALYALWSLSRHAALLYRGGYFSGEQAGEVLESAVLALCSQLKDDAVALVDVIAPPDFVLDSPIGRADGELYKNLWGAVLQESKVLERASWWPEFSVNKPVIGSLKSKL
3809	A	117	830	CFGIMERVGCTLTTTYAHPRPTPTNFLPAISTMASSYRDRFPHSNLTHSLSLPWRPSTYYKVASNSPSVAPYCTRSQRVSENTMLPFVSNRTTFFTRYTPDDWYRSNLTNYQESNTSRHNSEKLRVDTSRLIQDKYQQTRKTQADTTQNLGERVNDIGFWKSEIIHELDEMIGETNALTDVKKRLERALMETEAPLQVARECLFHREKRMGIDLVHDEVEAQLLTVNVGEMHQSQAA
3810	A	3	518	VIQELEGGSGADLGEHSCRPASQPRFPRPAEARSHPATRRPASGPAMGKTNSKLAPEVLEDLVQNTEFSEQELKQWYKGFLKDCPSGILNLEEFQQLYIKFFPYGDASKFAQHAFRTFDKNGDGTIDFREFICALSVTSRGSFEQKLNWAFEMYDLDGDGRITRLEMLEIIE
3811	A	81	1147	GCGYGCSGAGGAAIGEPMAKWGEGDPRWIVEERADATNVNNWHWTERDASNWSTDKLKTLFLAVQVQNEEGKCEVTEVSKLDGEASINNRKGKLIFFYEWSVKLNWTGTSKSGVQYKGHVEIPNLSDENSVDEVEISVSLAKDEPDTNLVALMKEEGVKLLREAMGIYISTLKTEFTQGMILPTMNGESVDPVGQPALKTEERKAKPAPSKTQARPVGVKIPTCKITLKETFLTSPEELYRVFTTQELVQAFTHAPATLEADRGGKFHMVDGNVSGEFTDLVPEKHIVMKWRFKSWPEGHFATITLTFIDKNGETELCMEGRGIPAPEEERTRQGWQRYYFEGIKQTFGYGARLF
3812	A	20	558	PCGTAASTHAYDRRAKCRQQQQQQNGGQNKVRPAKKKTSPAREVSSESGTSGQFTPPSSTSVPTIAS

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SSAPVSIWSPASISPLSDPLSTSSSCMQRSYPMTYT QASGYSQGYAGSTSYFGGMDCGSYLTPMHHQL PGPGATLSPMGTNAVTSHLNQSPASLSTQGYGAS KLWGFNFNH
3813	A	1	1016	CTEPPRRSTRTPAALASLRPYTDYVVVSDQILQES EDFFTLIESHEGKPLKLMVYNSKSDSCREVTVTP NAAWGGEGSLGCGIGYGYLHRIPTQPPSYHKKPP GTPPPSALPLGAPPPDALPPGPTPEDSPSLETGSRQ SDYMEALLQAPGSSMEDPLPGPGSPSHSAPDPDG LPHFMETPLQPPPPVQRVMDPGFLDVSGISLLDN SNASVWPSLPSSTELTTTAVSTSGPEDICSSSSSHE RGGEATWSGSEFEVSFLDSPGAQAQADHLPQLT LPDSLTSAASPEDGLSAELLEAQAEEEPASTEGLD TGTEAEGLDSQAQISTTE*HPGL*QGP
3814	A	2	884	VFWQVRNAGSSPLSAACPLFRTPAPQPCGSWGR CCIPHASTGCRPMAERGELDLTGAKQNTGVWLV KVPKYLSQQWAKASGRGEVGKLRIAKTQGRTE VSFTLNEDLANIHDIGGKPASVSAPREHPFVLQSV GGQTLTVFTESSSDKLSLEGIVVQRAECRPAASE NYMRLKRLQIEESSKPVRLSQQLDKVVTTNYKP VANHQYNIEYERKKKEDGKRARADKQHVLDML FSAFEKHQYYNLKDLVDITKQPVVYLKEILKEIG VQNVKGIHKNTWELKPEYRHYQGEEKSD
3815	A	17	411	NIGDWEDIGKSPERIIQYYGPATWAQDGSRGYCT PIYMLNHIIRLQAVLEIIMNERANALDLLAQQTTK MRNANYQNRLALDYLLAHEGGV*GKFSLTNCC LEIDDNGKAIMEITARMRKLAHIPVQTWER
3816	A	3	1172	SHWQRRDRRCVRNMAERGRKRPCGPGEHGQRI EWRKWKQQKKEEKKKWKDLKLMKKLERQRAQ EEQAKRLEEEEAAAEKEDRGRPYTLSVALPGSIL DNAQSPELRTYLAGQIARACAIFCVDEIVVFDEE GQDAKTVEGEFTGVGKKGQACVQLARILQYLEC PQYLRKAFFPKHQDLQFAGLLNPLDSPHHMRQD EESEFREGVVVDRPTRPGHGSFVNCGMKKEVKI DKNLEPGLRVTVRLNQQQHPDCKTYHGKVVSS QDPRTKAGLYWGYTVRLASCLSAVFAEAPFQDG YDLTIGTSERGSDVASAQLPNFRHALVVFGGLQG LEAGADADPNLEVAEPSVLFDLYVNTCPGQGSR TIRTEEAILISLAALQPGLIQAGARHT
3817	A	246	1197	FLSAGMSNFTHYAYLLMIESLMLGKVPPHVPSH HFIFHDDGSARQKGESDYKVIIQQWFSKSGPWTT SSNVTWGLLELQQSISESAVLTIPPGDSGAGSNLI TMFLRNRKETDLCSGRSKVNRGWNSGRCKQRG KTEQPGEPLEHVYVTIKHAVALESRHQKGELQC LIKMCIPLSKPLQMFFSPPHWEAWLQRVQQLAK NTRYFRQRLQEMGFIIYGNENASVVPLLLYMPG KVAAFARHMLEKKIGVVVVGFPATPLAEARARF CVSAAHTREMLDTVLEALDEMGDLLQLKYSRH KKSARPELYDETSFELED
3818	A	215	789	NPQSSSSEGSSEIFQVNGHNRLLVQRSEVTQAPG QYTVDVEGHGCTFIQATLKYNVLLPKKASGFSLS LEIVKNYSSTAFDLTVTLKYTGIRNKSSMVVIDV KMLSGFTPTMSSIEELENKGQVMKTEVKNDHVL FYLENVFGRADSFTFSVEQSNLVFNIQPAPGMVY DYYEKEEYALAFYHINSSSVSE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3819	A	1	1483	RIPDSIISRGVQGLPRDTASLSTTPSESPRAQATSR LSTASCPTPKVQSRCSSKENILRASHSAVDITKVA RRHRMSPFPLTSMDKAFITVLEMTPVLGTEIINYR DGMGRVLAQDVYAKDNLPPFPASVKDGYAVRA ADGPGDRFIIGESQAGEQPTQTVMPGQVMRVTT GAPIPCGADAVVQVEDTELIRESDDGTEELEVRIL VQARPGQDIRPIGHDIKRGECVLAKGTHMGPSEI GLLATVGVTEVEVNKFPVVAVMSTGNELLNPED DLLPGKIRDSNRSTLLATIQEHGYPTINLGIVGDN PDDLLNALNEGISRADVIITSGGVSMGEKDYLKQ VLDIDLHAQIHFGRVFMKPGLPTTFATLDIDGVR KIIFALPGNPVSAVVTCNLFVVPALRKMQGILDP RPTIIKARLSCDVKLDPRPEYHRCILTWHHQEPLP WAQSTGNQMSSRLMSMRSANGLLMLPPKTEQY VELHKGEVVDVMVIGRL
3820	A	2216	487	PQEPALKSEFSQVASNTIPLPLPQPNTCKDNGPCK QVCSTVGGSAICSCFPGYAIMADGVSCEDQDECL MGAHDCSRRQFCVNTLGSFYCVNHTVLCADGYI LNAHRKCVDINECVTDLHTCSRGEHCVNTLGSF HCYKALTCEPGYALKDGECEDVDECAMGTHTC QPGFLCQNTKGSFYCQARQRCMDGFLQDPEGNC VDINECTSLSEPCRPGFSCINTVGSYTCQRNPLIC ARGYHASDDGTKCVDVNECETGVHRCGEGQVC HNLPGSYRCDCKAGFQRDAFGRGCIDVNECWAS PGRLCQHTCENTLGSYRCSCASGFLLAADGKRC EDVNECEAQRCSQECANIYGSYQCYCRQGYQLA EDGHTCTDIDECAQGAGILCTFRCLNVPGSYQCA CPEQGYTMTANGRSCKDVDECALGTHNCSEAET CHNIQGSFRCLRFECPPNYVQVSKTKCERTTCHD FLECQNSPARITHYQLNFQTGLLVPAHIFRIGPAP AFTGDTIALNIIKGNEEGYFGTRRLNAYTGVVYL QRAVLEPRDFALDVEMKLWRQGSVTTFLAKMHI FFTTFAL
3821	A	2216	487	PQEPALKSEFSQVASNTIPLPLPQPNTCKDNGPCK QVCSTVGGSAICSCFPGYAIMADGVSCEDQDECL MGAHDCSRRQFCVNTLGSFYCVNHTVLCADGYI LNAHRKCVDINECVTDLHTCSRGEHCVNTLGSF HCYKALTCEPGYALKDGECEDVDECAMGTHTC QPGFLCQNTKGSFYCQARQRCMDGFLQDPEGNC VDINECTSLSEPCRPGFSCINTVGSYTCQRNPLIC ARGYHASDDGTKCVDVNECETGVHRCGEGQVC HNLPGSYRCDCKAGFQRDAFGRGCIDVNECWAS PGRLCQHTCENTLGSYRCSCASGFLLAADGKRC EDVNECEAQRCSQECANIYGSYQCYCRQGYQLA EDGHTCTDIDECAQGAGILCTFRCLNVPGSYQCA CPEQGYTMTANGRSCKDVDECALGTHNCSEAET CHNIQGSFRCLRFECPPNYVQVSKTKCERTTCHD FLECQNSPARITHYQLNFQTGLLVPAHIFRIGPAP AFTGDTIALNIIKGNEEGYFGTRRLNAYTGVVYL QRAVLEPRDFALDVEMKLWRQGSVTTFLAKMHI FFTTFAL
3822	A	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWDLAS LRCTLGAFCECDFRPDLPGLECDLAQHLAGQHL AKALVVKALKAFVRDPAPTKPLVLSLHGWTGTG KSYVSSLLAHYLFQGGLRSPRVHHFSPVLHFPHP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SHIERYKKDLKSWVQGNLTACGRSLFLFDEMDKMPPGLMEVLRPFLGSSWVVYGTNYRKAIFIFISNTGGEQINQVALEAWRSRRDREEILLQELEPVISRAVLDNPHHGFSNSGIMEERLLDAVVPFLPLQRHHVRHCVLNELAQLGLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCKTVASRIAFFL
3823	A	1	3174	YGCEKTTEGRIPLKNIYRLFSADRKRVETALEACSLPSSRNDSIPQEDFTPEVYRVFLNNLCPRPEIDNIFSEFGAKSKPYLTVDQMMDFINLKQRDPRLNEILYPPLKQEQVQVLIEKYEPNNSLARKGQISVDGFMRYLSGEENGVVSPEKLDLNEDMSQPLSHYFINSSHNTYLTAGQLAGNSSVEMYRQVLLSGCRCVELDCWKGRTAEEEPVITHGFTMTTEISFKEVIEAIAECAFKTSPFPILLSFENHVDSPKQQAKMAEYCRLIFGDALLMEPLEKYPLESGVPLPSPMDLMYKILVKNKKKSHKSSEGSGKKKLSEQASNTYSDSSSMFEPSSPGAGEADTESDDDDDDDDCKKSSMDEGTAGSEAMATEEMSNLVNYIQPVKFESFEISKKRNKSFEMSSFVETKGLEQLTKSPVEFVEYNKMQLSRIYPKGTRVDSSNYMPQLFWNAGCQMVALNFQTMDLAMQINMGMYEYNGKSGYRLKPEFMRRPDKHFDPFTEGIVDGIVANTLSVKIISGQFLSDKKVGTYVEVDMFGLPVDTRRKAFKTKTSQGNAVNPVWEEEPIVFKKVVLPTLACLRIAVYEEGGKFIGHRILPVQAIRPGYHYICLRNERNQPLTLPAVFVYIEVKDYVPDTYADVIEALSNPIRYVNLMEQRAKQLAALTLEDEEEVKKEADPGETPSEAPSEARTTPAENGVNHTTTLTPKPPSQALHSQPAPGSVKAPAKTEDLIQSVLTEVEAQTIEELKQQKSFVKLQKKHYKEMKDLVKRHHKKTTDLIKEHTTKYNEIQNDYLRRRAALEKSAKKDSKKKSEPSSPDHGSSTIEQDLAALDAEMTQKLIDLKDKQQQQLLNLRQEQYYSEKYQKREHIKLLIQKLTDVAEECQNNQLKKLKEICEKEKKELKKKMDKKRQEKITEAKSKDKSQMEEEKTEMIRSYIQEVVQYIKRLEEAQSKRQEKLVEKHKEIRQQILDEKPKLQVELEQEYQDKFKRLPLEILEFVQEAMKGKISEDSNHGSAPLSLSSDPGKVNHKTPSSEELGGDIPGKEFDTPL
3824	A	1	426	ILHWFVHRWSGRNNREKIGVHVGFEEILNMEPYCCRETLKSLRPECFIYDLSAVVMHHGKGFGSGHYTAYCYNSEGGFWVHCNDSKLSMCTMDEVCKAQAYILFYTQRVTENGHSKLLPPELLLGSQHPNEDADTSSNEILS
3825	A	3	364	GIRAKFPNKIPVVVERYPRETFLPPLDKTKFLVPQELTMTQFLSIIRSRMVLRATEAFYLLVNNKSLVSMSATMAEIYRDYKDEDGFVYMTYASQETFGCLESAAPRDGSSLEDRPLHPL
3826	A	1	1237	PEKKFERECREAEKAQQSYERLDNDTNATKADVEKAKQQLNLRTHMADENKNEYAAQLQNFNGEQHKHFYVVIPQIYKQLQEMDERRTIKLSECYRGFADSERKVIPIISKCLEGMILAAKSVDERRDSQMVVDSFKSGFEPPGDFPFEDYSQHIYRTISDGTISASKQESGKMDAKTTVGKAKGKLWLFGKKPKGPALEDFSHLPPEQRRKKLQQRIDELNRELQKESDQKDALNKMKDVYEKNPQMGDPGSLQPKLAETMNNIDR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LRMEIHKNEAWLSEVEGKTGGRGDRRHSSDINHLVTQGRESPEGSYTDDANQEVRGPPQQHGHHNEFDDEFEDDDPLPAIGHCKAIYPFDGHNEGTLAMKEGEVLYIIEEDKGDGWTRARRQNGEEGYVPTSYIDVTLEKNSKGS
3827	A	2	1584	INPVSSAVNGEAHSSHETRGQNSNALPSVLLELLSQSCLIPAMSSYLRNDSVLDMARHVPLYRALLELLRAIASCAAMVPLLLPLSTENGEEEEEQSECQTSVGTLLAKMKTCVDTYTNRLRSKRENVKTGVKPDASDQEPEGLTLLVPDIQKTAEIVYAATTSLRQANQEKKLGEYSKKAAMKPKPLSVLKSLEEKYVAVMKKLQFDTFEMVSEDEDGKLGFKVNYHYMSQVKNANDANSAARARRLAQEAVTLSTSLPLSSSSSVFVRCDEERLDIMKVLITGPADTPYANGCFEFDVYFPQDYPSSPPLVNLETTGGHSVRFNPNLYNDGKVCLSILNTWHGRPEEKWNPQTSSFLQVLVSVQSLILVAEPYFNEPGYERSRGTPSGTQSSREYDGNIRQATVKWAMLEQIRNPSPCFKEVIHKHFYLKRVEIMAQCEEWIADIQQYSSDKRVGRTMSHHAAALKRHTAQLREELLKLPCPEGLDPDTDDAPEVCRATTGAEETLMHDQVKPSSSKELPSDFQL
3828	A	1415	845	PRVPATLVSLDPWHCFPTAGRLAGSTWVPPACTLQLGPSSEHELDNHRAPLLSLPSQESLSFTPWYLVACKPLFHIFCPLFACFMQEGKVQYLFLHLSHMRLLNYYFFPFLAPESLMQALEDLDYLAALDNDGNLSEFGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIAAMVTGILNDYSFSFFANLH
3829	A	199	683	VDHTPVLSKPQCFSSVKWGATLSARSQKTSGIGRLMVHVIEATELKACKPNGKSNPYCEISMGSQSYTTRTIQDTLNPKWNFNCQFFIKDLYQDVLCLTLFDRDQFSPDDFLGRTEIPVAKIRTEQESKGPMTRRLLLHEVPTGEVWVRFDLQLFEQKTLL
3830	A	1747	404	RKMMEESGIETTPPGTPPPNPAGLAATAMSSTPVPLAATSSFSSPNVSSMESFPPLAYSTPQPPLPPVRPSAPLPFVPPPAVPSVPPLVTSMPPPVSPSTAAAFGNPPVSHFPPSTSAPNTLLPAPPSGPPISGFSVGSTYDITRGHAGRAPQTPLMPSFSAPSGTGLLPTPITQQASLTSLAQGTGTTSAITFPEEQEDPRITRGQDEASAGGIWGFIKGVAGNPMVKSVLDKTKHSVESMITTLDPGMAPYIKSGGELDIVVTSNKEVKVAAVRDAFQEVFGLAVVVGEAGQSNIAPQPVGYAAGLKGAQERIDSLRRTGVIHEKQTAVSVENFIAELLPDKWFDIGCLVVEDPVHGIHLETFTQATPVPLEFVQQAQSLTPQDYNLRWSGLLVTVGEVLEKSLLNVSRTDWHMAFTGMSRRQMIYSAARAIAGMYKQRLPPRTV
3831	A	5	674	FWTRSAWHEGLQQMKANDPSLQEVNLYNIKNIPIPTLREFAKALETNTHVKKFSLAATRSNDPVAIAFADMLKVNTTLTSLNIESHFITGTGILALVEALKENDTLTEIKIDNQRQQLGTAVEMEIAQMLEENSRILKFGYQFTKQGPRTRVAAAITKNNDLAWQKDTQEQTSIWQVVSQSIAGFNPQFEVQGQNARSWMEELGKAFHQFVRRELKQTEGKLP
3832	A	164	782	EPWVPMDVAESPERDPHSPEDEEQPQGLSDDDILRDSGSDQDLDGAGVRASDLEDEESAARGPSQEE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EDNHSDEEDRASEPKSQDQDSEVNELSRGPTSSP CEEEGDEGEEDRTSDLRDEASSVTRELDEHELDY DEEVPEEPAPAVQEDEAEKAGAEDDEEKGEGTP REEGKAGVQSVGEKESLEAAKEKKKEDDDGEID DEEMY
3833	A	122	1676	SQPPHFTQKMNENKDTDSKKSEEYEDDFEKDLE WLINENEKSDASIIEMACEKEENINQDLKENETV MEHTKRHSDPDKSLQDEVSPRRNDIISVPGIQPLD PISDSDSENSFQESKLESQKDLEEEEDEEVRRYIM EKIVQANKLLQNQEPVNDKRERKLKFKDQLVDL EVPPLEDTTTSKNYFENERNMFGKLSQLCISNDF GQEDVLLSLTNGSCEENKDRTILVERDGKFELLN LQDIASQGFLPPINNANSTENDPQQLLPRSSNSSV SGTKKEDSTAKIHAVTHSSTGEPLAYIAQPPLNR KTCPSSAVNSDRSKGNGKSNHRTQSAHISPVTST YCLSPRQKELQKQLEEKREKLKREEERRKIEEEK EKKRENDIVFKAWLQKKREQVLEMRRIQRAKEI EDMNSRQENRDPQQAFRLWLKKKHEEQMKERQ TEELRKQEECLFFLKGTEGRERAFKQWLRRKRM EKMAEQQAVRERTRQLRLEAKRSKQLQHHLYM SEAKPFRFTDHYN
3834	A	575	774	RSRTEELSNSGILKAMSKDLVTFGDVAVNFSQEE WEWLNPAQRNLYRKVMLENYRSLVSLGKDMSP
3835	A	2	100	ASDFYLRYYVGHKGKFGHEFLEFEFRPDGVYV
3836	A	91	749	RPTPGHGDFWMQPLTKDAGMSLSSVTLASALQV RGEALSEEEIWSLLFLAAEQLLEDLRNDSSDYVV CPWSALLSAAGSLSFQGRVSHIEAAPFKAPELLQ GQSEDEQPDASQMHVYSLGMTLYWSAGFHVPP HQPLQLCEPLHSILLTMCEDQPHRRCTLQSVLEA CRVHEKEVSVYPAPAGLHIRRLVGLVLGTISEVS REPCFSSSSCWSCVAIKI
3837	A	3	1214	SLGCTNSARGKGQDDEVRTLMANGAPFTTDWFS KLRVSCGYIGDNCKNGADVNAKDMLKMTALH WATERHHRDVVELLIKYGADVHAFSKFDKSAFD IALEKNNAEILVILQEAMQNQVNVNPERANPVTD PVSMAAPFIFTSGEVVNLASLISSTNTKTTSGDPH ASTVQFSNSTTSVLATLAALAEASVPLSNSHRAT ANTEEIIEGNSVDSSIQQVMGSGGQRVITIVTDGV PLGNIQTSIPTGGIGHPFIVTVQDGQQVLTVPAGK VAEETVIKEEEEEKLPLTKKPRIGEKTNSVEESKE GNERELLQQQLQEANRRAQEYRHQLLKKEQEAE QYRLKLEAIARQQPNGVDFTMVEEVAEVDAVV VTEGELEERETKVTGSAGATGPPTRVSMATVSS
3838	A	1	1332	MIEDNKENKDHSLERGRASLIFSLKNEVGGLIKA LKIFQEKHVNLLHIESRKSKRRNSEFEIFVDCDIN REQLNDIFHLLKSHTNVLSVNLPDNFTLKEDGME TVPWFPKKISDLDHCANRVLMYGSELDADHPGF KDNVYRKRRKYFADLAMNYKHGDPIPKVEFTEE EIKTWGTVFQELNKLYPTHACREYLKNLPLLSKY CGYREDNIPQLEDVSNFLKERTGFSIRPVAGYLSP RDFLSGLAFRVFHCTQYVRHSSDPFYTPEPDTCH ELLGHVPLLAEPSFAQFSQEIGLASLGASEEAVQ KLATCYFFTVEFGLCKQDGQLRVFGAGLLSSISE LKHALSGHAKVKPFDPKITCKQECLITTFQDVYF VSESFEDAKEKMREFTKTIKRPFGVKYNPYTRSI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				QILKDTKSITSAMNELQHDLDVVSDALAKVSRKPSI
3839	A	3093	520	MVNFTVDQIRAIMDKKANIRNMSVIAHVDHGKSTLTDSLVCKAGIIASARAGETRFTDTRKDEQERCITIKSTAISLFYELSENDLNFIKQSKDGAGFLINLIDSPGHVDFSSEVTAALRVTDGALVVVDCVSGVCVQTETVLRQAIAERIKPVLMMNKMDRALLELQLEPEELYQTFQRIVENVNVIISTYGEGESGPMGNIMIDPVLGTVGFGSGLHGWAFTLKQFAEMYVAKFAAKGEGQLGPAERAKKVEDMMKKLWGDRYFDPANGKFSKSATSPEGKKLPRTFCQLILDPIFKVFDAIMNFKKEETAKLIEKLDIKLDSEDKDKEGKPLLKAVMRRWLPAGDALLQMITIHLPSPVTAQKYRCELLYEGPPDDEAAMGIKSCDPKGPLMMYISKMVPTSDKGRFYAFGRVFSGLVSTGLKVRIMGPNYTPGKKEDLYLKPIQRTILMMGRYVEPIEDVPCGNIVGLVGVDQFLVKTGTITTFEHAHNMRVMKFSVSPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCIIEESGEHIIAGAGELHLEICLKDLEEDHACIPIKKSDPVVSYRETVSEESNVLCLSKSPNKHNRLYMKARPFPDGLAEDIDKGEVSARQELKQRARYLAEKYEWDVAEARKIWCFGPDGTGPNILTDITKGVQYLNEIKDSVVAGFQWATKEGALCEENMRGVRFDVHDVTLHADAIHRGGGQIIPTARRCLYASVLTAQPRLMEPIYLVEIQCPEQVVGGIYGVLNRKRGHVFEESQVAGTPMFVVKAYLPVNESFGFTADLRSNTGGQAFPQCVFDHWQILPGDPFDNSSRPSQVVAETRKRKGLKEGIPALDNFLDKL
3840	A	2	753	SSTRSRDFCCSEAIQGSLTRRERRASGVRTRRSQGSSAMASKILLNVQEEVTCPICLELLTEPLSLDCGHSLCRACITVSNKEAVTSMGGKSSCPVCGISYSFEHLQANQHLANIVERLKEVKLSPDNGKKRDLCDHHGEKLLLFCKEDRKVICWLCERSQEHRGHHTVLTEEVFKECQEKLQAVLKRLKKEEEEAEKLEADIREEKTSWKYQVQTERQRIQTEFDQLRSILNNEEQRELQRLEEEEKKT
3841	A	2	405	GKAFSCFTYLSQHRRTHMAEKPYECKTCKKAFSHFGNLKVHERIHTGEKPYECKECRKAFSWLTCLLRHERIHTGKKSYECQQCGKAFTRSRFLRGHEKTHTGEKMHECKECGKALSSLSSLHRHKRTHWRDTL
3842	A	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDFTSVVSEIMMYILLVFLTLWLLIEMIYCYRKVSKAEEAAQENA
3843	A	3	1175	APIRNSRIDDFVRRVESKATSARCGLWGSGPRRRPASGMFRGLSSWLGLQQPVAGGGQPNGDAPPEQPSETVAESAEEELQQAGDQELLHQAKDFGNYLFNFASAATKKITESVAETAQTIKKSVEEGKIDGIIDKTIIGDFQKEQKKFVEEQHTKKSEAAVPPWVDTNDEETIQQQILALSADKRNFLRDPPAGVQFNFDFDQMYPVALVMLQEDELLSKMRFALVPKLVKEEVFWRNYFYRVSLIKQSAQLTALAAQQQAAGKEEKSNGREQDLPLAEAVRPKTPPVVIKSQLKTQEDEEEISTSPGVSEFVSDAFDACNLNQEDLRKEMEQLVLDKKQEETAVLEEDSADWEKELQQELQEYEV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				VTESEKRDENWDKEIEKMLQEEN
3844	A	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDPLLL RCKLLPSALQKMALGMFFGFTSVIVAGVLEMER LHYIHHNETVSQQIGEVLYNAAPLSIWWQIPQYL LIGISEIFASIPGLEFAYSEAPRSMQGAIMGIFFCLS GVGSLLGSSLVALLSLPGGWLHCPKDFGNINNCR MDLYFFLLAGIQAVTALLFVWIAGRYERASQGP ASHSRFSRDRG
3845	A	3	1934	PEDSAPQYSRLFPNASQHITPSYNYAPNPDKHWI MRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDS METIYNMLVETGELDNTYIVYTADHGYHIGQFG LVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIV LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERP VNRFHLKKKMRVWRDSFLVERGKLLHKRDNDK VDAQEENFLPKYQRVKDLCQRAEYQTACEQLG QKWQCVEDATGKLKLHKCKGPMRLGGSRALSN LVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFK KKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAA QPRNLTKRHWPGAPEDQDDKDGGDFSGTGGLP DYSAANPIKVTHRCYILENDTVQCDLDLYKSLQ AWKDHKLHIDHEIETLQNKIKNLREVRGHLKKK RPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGL QEKDKVWLLREQKRKKKLRKLLKRLQNNDTCS MPGLTCFTHDNQHWQTAPFWTLGPFCACTSAN NNTYWCMRTINETHNFLFCEFATGFLEYFDLNT DPYQLMNAVNTLDRDVLNQLHVQLMELRSCKG YKQCNPRTRNMDLGLKDGGSYEQYRQFQRRKW PEMKRPSSKSLGQLWEGWEG
3846	A	3	1934	PEDSAPQYSRLFPNASQHITPSYNYAPNPDKHWI MRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDS METIYNMLVETGELDNTYIVYTADHGYHIGQFG LVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIV LNIDLAPTILDIAGLDIPADMDGKSILKLLDTERP VNRFHLKKKMRVWRDSFLVERGKLLHKRDNDK VDAQEENFLPKYQRVKDLCQRAEYQTACEQLG QKWQCVEDATGKLKLHKCKGPMRLGGSRALSN LVPKYYGQGSEACTCDSGDYKLSLAGRRKKLFK KKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAA QPRNLTKRHWPGAPEDQDDKDGGDFSGTGGLP DYSAANPIKVTHRCYILENDTVQCDLDLYKSLQ AWKDHKLHIDHEIETLQNKIKNLREVRGHLKKK RPEECDCHKISYHTQHKGRLKHRGSSLHPFRKGL QEKDKVWLLREQKRKKKLRKLLKRLQNNDTCS MPGLTCFTHDNQHWQTAPFWTLGPFCACTSAN NNTYWCMRTINETHNFLFCEFATGFLEYFDLNT DPYQLMNAVNTLDRDVLNQLHVQLMELRSCKG YKQCNPRTRNMDLGLKDGGSYEQYRQFQRRKW PEMKRPSSKSLGQLWEGWEG
3847	A	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITLPPP FQHPDLSPLLRYSFETMAPTGLSSLTVNSTAVPTT PAAFKSLNLPLQITLSAIMIFILFVSFLGNLVVCLM VYQKAAMRSAINILLASLAFADMLLAVLNMPFA LVTILTTRWIFGKFFCRVSAMFFWLFVIEGVAILL IISIDRFLIIVQRQDKLNPYRAKVLIAVSWATSFCV AFPLAVGNPDLQIPSRAPQCVFGYTTNPGYQAYV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				ILISLISFFIPFLVILYSFMGILNTLRHNALRIHSYPE GICLSQASKLGLMGLQRPFQMSIDMGFKTRAFTT ILILFAVFIVCWAPFTTYSLVATFSKHFYYQHNFF EISTWLLWLCYLKSALNPLIYYWRIKKFHDACLD MMPKSFKFLPQLPGHTKRRIRPSAVYVCGEHRT VV
3848	A	3	2827	SSAVAARRRRSWASLVLAFLGVCLGITLAVDRS NFKTCEESSFCKRQRSIRPGLSPYRALLDSLQLGP DSLTVHLIHEVTKVLLVLELQGLQKNMTRFRIDE LEPRRPRYRVPDVLVADPPIARLSVSGRDENSVE LTMAEGPYKIILTARPFRLDLLEDRSLLLSVNARG LLEFEHQRAPRVSQGSKDPAEGDGAQPEETPRD GDKPEETQGKAEKDEPGAWEETFKTHSDSKPYG PMSVGLDFSLPGMEHVYGIPEHADNLRLKVTEG GEPYRLYNLDVFQYELYNPMALYGSVPVLLAHN PHRDLGIFWLNAAETWVDISSNTAGKTLFGKMM DYLQGSGETPQTDVRWMSETGIIDVFLLLGPSISD VFRQYASLTGTQALPPLFSLGYHQSRWNYRDEA DVLEVDQGFDDHNLPCDVIWLDIEHADGKRYFT WDPSRFPQPRTMLERLASKRRKLVAIVDPHIKVD SGYRVHEELRNLGLYVKTRDGSDYEGWCWPGS AGYPDFTNPTMRAWWANMFSYDNYEGSAPNLF VWNDMNEPSVFNGPEVTMLKDAQHYGGWEHR DVHNIYGLYVHMATADGLRQRSGGMERPFVLA RAFFAGSQRFGAVWTGDNTAEWDHLKISIPMCL SLGLVGLSFCGADVGGFFKNPEPELLVRWYQMG AYQPFFRAHAHLDTGRREPWLLPSQHNDIIRDAL GQRYSLLPFWYTLLYQAHREGIPVMRPLWVQYP QDVTTFNIDDQYLLGDALLVHPVSDSGAHGVQV YLPGQGEVWYDIQSYQKHHGPQTLYLPVTLSSIP VFQRGGTIVPRWMRVRRSSECMKDDPITLFVALS PQGTAQGELFLDDGHTFNYQTRQEFLLRRFSFSG NTLVSSSADPEGHFETPIWIERVVIIGAGKPAAVV LQTKGSPESRLSFQHDPETSVLVLRKPGINVASD WSIHLR
3849	A	1	1717	RARNARGCWGVCRSGFSSAVCGAARMEQVAEG ARVTAVPVSAADSTEELAEVEEGVGVVGEDNDA AARGAEAFGDSEEDGEDVFEVEKILDMKTEGGK VLYKVRWKGYTSDDDTWEPEIHLEDCKEVLLEF RKKIAENKAKAVRKDIQRLSLNNDIFEANSDSDQ QSETKEDTSPKKKKKKLRQREEKSPDDLKKKKA KAGKLKDKSKPDLESSLESLVFDLRTKKRISEAK EELKESKKPKKDEVKETKELKKVKKGEIRDLKT KTREDPKENRKTKKEKFVESQVESESSVLNDSPF PEDDSEGLHSDSREEKQNTKSARERAGQDMGLE HGFEKPLDSAMSAEEDTDVRGRRKKKTPRKAED TRENRKLENKNAFLEKKTVPKKQRNQDRSKSAA ELEKLMPVSAQTPKGRRLSGEERGLWSTDSAEE DKETKRNESKKPKKDEVKETKELKKVKKGEIRD LKTKTREDPKENRKTKKEKFVESQVESESSVLND SPFPEDDSEGLHSDSREEKQNTKSARERAGQDM GLEHGFEKPLDSAMSAEEDTDVRGRRKKKTPRK AEDTRENRKLENKNAFLEKKTVPKKQRNQDRSK SAAELEKLMPVSAQTPKGRRLSGEERGLWSTDS AEEDKETKRNESKKPKKDEVKETKELKKVKKGE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				IRDLKTKTREDPKENRKTKKEKFVESQVESESSV LNDSPFPED/RQ*RATFRQQREEKSPDDLKKKKA KAGKLKDKSKPDLESSLESLVFDLRTKKRISEAK EELKESKKPK
3850	A	1113	3975	PAAAAAAAAAAAAAGRGPSFTPCFSPSLAVEPS RRTRLGSDPAQAMAGNVKKSSGAGGGSGSGGS GSGGLIGLMKDAFQPHHHHHHHLSPHPPGTVDK KMVEKCWKLMDKVVRLCQNPKLALKNSPPYIL DLLPDTYQHLRTILSRYEGKMETLGENEYFRVF MENLMKKTKQTISLFKEGKERMYEENSQPRRNL TKLSLIFSHMLAELKGIFPSGLFQGDTFRITKADA AEFWRKAFGEKTIVPWKSFRQALHEVHPISSGLE AMALKSTIDLTCNDYISVFEFDIFTRLFQPWSSLL RNWNSLAVTHPGYMAFLTYDEVKARLQKFIHKP GSYIFRLSCTRLGQWAIGYVTADGNILQTIPHNKP LFQALIDGFREGFYLFPDGRNQNPDLTGLCEPTP QDHIKVTQEQYELYCEMGSTFQLCKICAENDKD VKIEPCGHLMCTSCLTSWQESEGQGCPFCRCEIK GTEPIVVDPFDPRGSGSLLRQGAEGAPSPNYDDD DDERADDTLFMMKELAGAKVERPPSPFSMAPQA SLPPVPPRLDLLPQRVCVPSSASALGTASKAASGS LHKDKPLPVPPTLRDLPPPPPPDRPYSVGAESRPQ RRPLPCTPGDCPSRDKLPPVPSSRLGDSWLPRPIP KVPVSAPSSSDPWTGRELTNRHSLPFSLPSQMEP RPDVPRLGSTFSLDTSMSMNSSPLVGPECDHPKI KPSSSANAIYSLAARPLPVPKLPPGEQCEGEEDTE YMTPSSRPLRPLDTSQSSRACDCDQQIDSCTYEA MYNIQSQAPSITESSTFGEGNLAAAHANTGPEES ENEDDGYDVPKPPVPAVLARRTLSDISNASSS/FG LFVLERDP*PQNVTEGSQVPERPPKPFPRRINSER KAGSCQQGSGPAASAATA\SPQLSSEIENLMSQG YSYQDIQKALVIAQNNIEMAKNILREFVSISSPAH VAT
3851	A	2	2781	GRVGSMDGAMGPRGLLLCMYLVSLLILQAMPA LGSATGRSKSSEKRQAVDTAVDGVFIRSLKVNC KVTSRFAHYVVTSQVVNTANEAREVAFDLEIPK TAFISDFAVTADGNAFIGDIKDKVTAWKQYRKA AISGENAGLVRASGRTMEQFTIHLTVNPQSKVTF QLTYEEVLKRNHMQYEIVIKVKPKQLVHHFEIDV DIFEPQGISKLDAQASFLPKELAAQTIKKSFSGKK GHVLFRPTVSQQQSCPTCSTSLLNGHFKVTYDVS RDKICDLLVANNHFAHFFAPQNLTNMNKNVVFV IDISGSMRGQKVKQTKEALLKILGDMQPGDYFD LVLFGTRVQSWKGSLVQASEANLQAAQDFVRGF SLDEATNLNGGLLRGIEILNQVQESLPELSNHASI LIMLTDGDPTEGVTDRSQILKNVRNAIRGRFPLY NLGFGHNVDFNFLEVMSMENNGRAQRIYEDHD ATQQLQGFYSQVAKPLLVDVDLQYPQDAVLALT QNHHKQYYEGSEIVVAGRIADNKQSSFKADVQA HGEGQEFSITCLVDEEEMKKLLRERGHMLENHV ERLWAYLTIQELLAKRMKVDREVRANLSSQALR MSLDYGFVTPLTSMSIRGMADQDGLKPTIDKPSE DSPPLEMLGPRRTFVLSALQPSPTHSSSNTQRLPD RVTGVDTDPHFIIHVPQKEDTLCFNINEEPGVILS LVQDPNTGFSVNGQLIGNKARSPGQHDGTYFGR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LGIANPATDFQLEVTPQNITLNPGFGGPVFSWRDQAVLRQDGVVVTINKKRNLVVSVDDGGTF\EVV\LHRVW\KGSS\VHQDFLGLLMCWDKSIGMSSPGRKGCWGQ\FFHPIRFLKVS*HPPPGSDPQKAQMPTMVVRNPPGLTVT\RGLQKDYSKDPWHGAEVSCWFI\HNNGA*I\TDCAYTDYI\VPDIF
3852	A	39	1735	TQVAEAGRGEGVVAGAETGRPQSAGMNLELLESFGQNYPEEADGTLDCISMALTCTFNRWGTLLAVGCNDGRIVIW\DF\LTRGIA*NKFSAHIHPVCSLCWSRDGHKLVSASTDNIVSQWDVLSGDCDQRFRFPSPILKVQYHPRDQNKVLVCPMKSAPVMLTLSDSKHVVLPVDDDSDLNVVASFDRRGEYIYTGNAKGKILVLKTDSQDLVASFRVTTGTSNTTAIKSIEFARKGSCFLINTADRIIRVYDGREILTCGRDGEPEPMQKLQDLVNRTPWKKCCFSGDGEYIVAGSARQHALYIWEKSIGNLVKILHGTRGELLLDVAWHPVRPIIASISSGVVSIWAQNQVENWSAFAPDFKELDENVEYEERESEFDIEDEDKSEPEQTGADAAEDEEVDVTSVDPIAAFCSSDEELEDSKALLYLPIAPEVEDPEENPYGPPPDAVQTSLMDEGASSEKKRQSSADGSQPPKKKPKTTNIELQGVPNDEVHPLLGVKGDGKSKKKQAGRPKGSKGKEKDSPFKPKLYKGDRGLPLEGSAKGKVQAELSQPLTAGGAISELL
3853	A	45	2603	PLLFTCGREVRARDPEKEGTIVVAGLKVQVQPRFLWILCFSMEETQGELTSSCGSKTMANVSLAFRDVSIDLSQEEWECLDAVQRDLYKDVMLENYSNLVSLDLEYKYITKNLLSEKNVCKIYLSQLQTGEKSKNTIHEDTIFRNGLQCKHEFERQERHQMGCVSQMLIQKQISHPLHPKIHAREKSYECKECRKAFRQQSYLIQHLRIHTGERPYKCMECGKAFCRVGDLRVHHTIHAGERPYECKECGKAFRLHYHLTEHQRIHSGVKPYECKECGKAFSRVRDLRVHQTIHAGERPYECKECGKAFRLHYQLTEHQRIHTGERPYECKVCGKTFRVQRHISQHQKIHTGVKPYKCNECGKAFSHGSYLVQHQKIHTGEKPYECKECGKSFSFHAELARHRRIHTGEKPYECRECGKAFRLQTELTRHHRTHTGEKPYECKECGKAFICGYQLTLHLRTHTGEIPYECKECGKTFSSRYHLTQHYRIHTGEKPYICNECGKAFRLQGELTRHHRIHTCEKPYECKECGKAFIHSNQFISHQRIHTSESTYICKECGKIFSRRYNLTQHFKIHTGEKPYICNECGKAFRFQTELTQHHRIHTGEKPYKCTECGKAFIRSTHLTQHHRIHTGEKPYECTECGKTFSRHYHLTQHHRGHTGEKPYICNECGNAFICSYRLTLHQRIHTGELPYECKECGKTFSRRYHLTQHFRLHTGEKPYSCKECGNAFRLQAELTRHHIVHTGEKPYKCKECGKAFSVNSELTRHHRIHTGEKPYQCKECGKAFIRSDQLTLHQ\KIILVR\NPMHNVKRIRWPLENAL*QRICNLRNFLFVTEHVGIPFTSCSQFIRNYFVC
3854	A	108	894	LQSCWVPGIPWPSVGWLSWLKDLPSCEIHSASLSAVLQGPQCSEMLWPKNLTSWDDSSSVSSGISDTIDNLSTDDINTSSSISSYANTPASSRKNLDVQTDAEKHSQVERNSLWSGDDVKKSDGGSDSGIKMEPGSKWRRNPSDVSDESDKSTSGKKNPVISQTGSWRRGMTAQVGITMPRTKASAPAGALKTPGTGKRPGL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				S\GPGAPTPAAPPQLARMAWAFSLSAASTPAVSPSTSPSAVEGSPATILPLASSPPPRTTP*LPLSELTV*RPQELVRGRGCLGPGAPTPAAPPQLARMAWAFSLSAASTPAVSPSTSPSAVEGSPATILPLASSPPPRTTP
3855	A	1	772	FRGGDGAPGVLKPGNPLPFPLPPLQYPPPSTLSHSDNLAMTSRSTARPNGQPQASKICQFKLVLLGESAVGKSSLVLRFVKGQFHEYQESTIGAAFLTQSVCLDDTTVKFEIWDTAGQERYHSLAPMYYRGAQAAIVVYDITNQETFARAKTWVKELQRQASP\SIVVGLAGNKADLANKRMVEYEEAQAYADDNSLLFMETSAKTAMNVNDLFL\AIA*EVAKRVNPQNLG\G\AAGRSRGVDLHEQS\QQNKSQCCSN
3856	A	2815	352	LGLEAAARPRPGGPAAMQDGNFLLSALQPEAGVCSLALPSDLQLDRRGAEGPEAERLRAARVQEQVRARLLQLGQQPRHNGAAEPEPEAETARGTSRGQYHTLQAGFSSRSQGLSGDKTSGFRPIAKPAYSPASWSSRSAVDLSCSRRLSSAHNGGSAFGAAGYGGAQPTPPMPTRPVSFHERGGVGSRADYDTLSLRSLRLGPGGLDDRYSLVSEQLEPAATSTYRAFAYERQASSSSSRAGGLDWPEATEVSPSRTIRAPAVRTLQRFQSSHRSRGVGGAVPGAVLEPVARAPSVRSLSLSLADSGHLPDVHGFNSYGSHRTLQRLSSGFDDIDLPSAVKYLMASDPNLQVLGAAYIQHKCYSDAAAKKQARSLQAVPRLVKLFNHANQEVQRHATGAMRNLIYDNADNKLALVEENGIFELLRTLREQDDELRKNVTGILWNLSSSDHLKDRLAKKTPLE\QLT\DLGV*APLSGAGGPP\LIQQNASEAEIFYNATGFPRNLSSASQATRQKMRECHGLVDALVTSINHALDAGKCEDKSVENAVCVLRNLSYRLYDEMPPSALQRLEGRGRRDLAGAPPGEVVGCFTPQSRRLRELPLAADALTFAEVSKDPKGLEWLWSPQIVGLYNRLLQRCELNRHTTEAAAGALQNITGG\DPRGPGGLSRLALEQERILNPLLDRVRTADHHQLRSLTGLIRNLSRNARNKDEMSTKVV\SHLI\EKLPGSVGEKSPPAEVLV\NI\IAVFNNLGWLASPI/ALARDLLYFDGLRKLIFIKKKRDSPDSEKSSRAASSLLANLWQYNKLHRDFRAKGYRKEDFLGP
3857	A	1034	204	VAVTLLSQLPSAIQRTAAWEMRAPLTFRVPLALDLIKPEHCTVNVDNSLSIPVIAAELVVRKPSEKGMQQKKKTKDLGFRAGKESKTEWRK*GLQDMASQMFALPLK*PVTAAFHDSSMPSSLLQIEMEQLFLEARLQ/PDSKSEARRNQCDSMLLRNQQLCSTCQEMKMVQPRTMKIPDDPKASFENCMSYRMSLHQPKFQTTPEPFHDDIPTENIHLQNL/PILGPRTAVFHGLLTEAYKTLKERQRSSLPRKEPIGKTTEAVSGRSSSPPRLPERK
3858	A	203	3469	SHQEIEQNSAMAPRKRGGRGISFIFCCFRNNDHPEITYRLRNDSNFALQTMEPALPMPPVEELDVMFSELVDELDLTDKHREAMFALPAEKKWQIYCSKKKDQEENKGATSWPEFYIDQLNSMAARKSLLALEKEEEEERSKTIESLKTALRTKPMRFVTRFIDLDGLSCILNFLKTMDYETSESRIHTSLIGCIKALMNNSQGRAHVLAHSESINVIAQSLSTENIKTKVAVLEILGAVCLVPGGHKKVLQAMLHYQKYASERTRFQTLIN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DLDKSTGRYRDEVSLKTAIMSFINAVLSQGAGVE SLDFRLHLRYE\FLMLGIHPVMDKLRKHENSTLD RHLDFFEMLRNEDELEFAKRFELVHIDTKSATQM FELTRKRLTHSEAYPHFMSILHHCLQMPYKRSGN TVQYWLLLDRIIQQIVIQNDKGQDPDSTPLENFNI KNVVRMLVNENEVKQWKEQAEKMRKEHNELQ QKLEKKERECDAKTQEKEEMMQTLNKMKEKLE KETTEHKQVKQQVADLTAQLHELSRRAVCASIP GGPSPGAPGGPFPSSVPGSLLPPPPPPPLPGGMLPP PPPPLPPGGPPPPPGPPPLGAIMPPPGAPMGLALK KKSIPQPTNALKSFNWSKLPENKLEGTVWTEIDD TKVFKILDLEDLERTFSAYQRQQDFFVNSNSKQK EADAIDDTLSSKLKVKELSVIDGRRAQNCNILLS RLKLSNDEIKRAILTMDEQEDLPKDMLEQLLKFV PEKSDIDLLEEHKHELDRMAKADRFLFEMSRINH YQQRLQSLYFKKKFAERVAEVKPKVEAIRSGSEE VFRSGALKQLLEVVLAFGNYMNKGQRGNAYGF KISSLNKIADTKSSIDKNITLLHYLITIVENKYPSV LNLNEELRDIPQAAKVNMTELDKEISTLRSGLKA VETELEYQKSQPPQPGDKFVSVVSQFITVASFSFS DVEDLLAEAKDLFTKAVKHFGEEAGKIQPDEFF GIFDQFLQAVSEAKQENENMRKKKEEEERRARM EAQLKEQRERERKMRKAKENSEESGEFDDLVSA LRSGEVFDKDLSKLKRNRKRITNQMTDSSRERPI TKLNF
3859	A	1279	141	RVEHLSEFLVDIKPSLTFDVIPLLDPYGPAGSDPS LEFLVVSEETYRGGMAINRFRLENDLEELALYQI QLLKDLRHTENEEDKVSSSSFRQRMLGNLLRPPY ERPELPTCLYVIGLTGISGSGKSSIAQRLKGLGAF VIDSDHLGHRAYAPGGPAYQPVVEAFGTDILHK DGIINRKVLGSRVFGNKKQLKILTDIMWPIIAKLA REEMDRAVAEGKRVCVIDAAVLLEAGWQNLVH EVWTAVIPETEAVRRIVERDGLSEAAAQSRLQSQ MSGQQLVEQSHVVLST\CGSRISPNARWRKPGPS CRSAFPRLIRPSTEKFSVGPDWLLELTSDPVVRRN GGLDAHPGSGPEVQAILCRTWPGLVDTGSLPNTL VFGQH
3860	A	1	3881	MGQKSVGASYVQIPLVPPLSRHPKGLGHEDRWS SYCLSSLAAQNICTSKLHCPAAPEHTDPSEPRGSV SCCSLLRGLSSGWSSPLLPAPVCNPNKAIFTVDA KTTEILVANDKACGLLGYSSQDLIGQKLTQFFLR SDSDVVEALSEEHMEADGHAAVVFGTVVDIISRS GEKIPVSVWMKRMRQERRLCCVVVLEPVERVST WVAFQSDGTVTSCDSLFAHLHGYVSGEDVAGQ HITDLIPSVQLPPSGQHIPKNLKIQRSVGRARDGT TFPLSLKLKSQPSSEEATTGEAAPVSGYRASVWV FCTISGLITLLPDGTIHGINHSFALTLFGYGKTELL GKNITFLIPGFYSYMDLAYNSSLQLPDLASCLDV GNESGCGERTLDPWQGQDPAEGGQDPRINVVLA GGHVVPRDEIRKLMESQDIFTGTQTELIAGGQLL SCLSPQPAPGVDNVPEGSLPVHGEQALPKDQQIT ALGREEPVAIESPGQDLLGESRSEPVDVKPFASCE DSEAPVPAEDGGSDAGMCGLCQKAQLERMGVS GPSGSDLWAGAAVAKPQAKGQLAGGSLLMHCP CYGSEWGLWWRSQDLAPSPSGMAGLSFGTPTLD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EPWLGVENDREELQTCLIKEQLSQLSLAGALDVP HAELVPTECQAVTAPVSSCDLGGRDLCGGCTGS SSACYALATDLPGGLEAVEAQEVDVNSFSWNLK ELFFSDQTDQTSSNCSCATSELRETPSSLAVGSDP DVGSLQEQGSCVLDDRELLLLTGTCVDLGQGRR FRESCVGHDPTEPLEVCLVSSEHYAASDRESPGH VPSTLDAGPEDTCPSAEEPRLNVQVTSTPVIVMR GAAGLQREIQEGAYSGSCYHRDGLRLSIQFEVRR VELQGPTPLFCCWLVKDLLHSQRDSAARTRLFL ASLPGSTHSTAAELTGPSLVEVLRARPWFEEPPK AVELEGLAACEGEYSQKYSTMSPLGSGAFGFVW TAVDKEKNKEVVVKFIKKEKVLEDCWIEDPKLG KVTLEIAILSRVEHANIIKVLDIFENQGFFQLVME KHGSGLDLFAFIDRHPRLDEPLASYIFRQVRAG\Q SRLVSAVGYLRLKDIIHRDIKDENIVIAEDFTIKLI DFGSAAYLERGKLFYTFCGTIEYCAPEVLMGNPY RGPELEMWSLGVTLYTLVFEENPFCELEETVEAA IHPPYLVSKELMSLVSGLLQPVPERRTTLEKLVT DPWVTQPVNLADYTWEEVFRVNKPESGVLSAAS LEMGNRSLSDVAQAQELCGGPVPGEAPNGQGCL HPGDPRLLTS
3861	A	1	3881	MGQKSVGASYVQIPLVPPLSRHPKGLGHEDRWS SYCLSSLAAQNICTSKLHCPAAPEHTDPSEPRGSV SCCSLLRGLSSGWSSPLLPAPVCNPNKAIFTVDA KTTEILVANDKACGLLGYSSQDLIGQKLTQFFLR SDSDVVEALSEEHMEADGHAAVVFGTVVDIISRS GEKIPVSVWMKRMRQERRLCCVVVLEPVERVST WVAFQSDGTVTSCDSLFAHLHGYVSGEDVAGQ HITDLIPSVQLPPSGQHIPKNLKIQRSVGRARDGT TFPLSLKLKSQPSSEEATTGEAAPVSGYRASVWV FCTISGLITLLPDGTIHGINHSFALTLFGYGKTELL GKNITFLIPGFYSYMDLAYNSSLQLPDLASCLDV GNESGCGERTLDPWQGQDPAEGGQDPRINVVLA GGHVVPRDEIRKLMESQDIFTGTQTELIAGGQLL SCLSPQPAPGVDNVPEGSLPVHGEQALPKDQQIT ALGREEPVAIESPGQDLLGESRSEPVDVKPFASCE DSEAPVPAEDGGSDAGMCGLCQKAQLERMGVS GPSGSDLWAGAAVAKPQAKGQLAGGSLLMHCP CYGSEWGLWWRSQDLAPSPSGMAGLSFGTPTLD EPWLGVENDREELQTCLIKEQLSQLSLAGALDVP HAELVPTECQAVTAPVSSCDLGGRDLCGGCTGS SSACYALATDLPGGLEAVEAQEVDVNSFSWNLK ELFFSDQTDQTSSNCSCATSELRETPSSLAVGSDP DVGSLQEQGSCVLDDRELLLLTGTCVDLGQGRR FRESCVGHDPTEPLEVCLVSSEHYAASDRESPGH VPSTLDAGPEDTCPSAEEPRLNVQVTSTPVIVMR GAAGLQREIQEGAYSGSCYHRDGLRLSIQFEVRR VELQGPTPLFCCWLVKDLLHSQRDSAARTRLFL ASLPGSTHSTAAELTGPSLVEVLRARPWFEEPPK AVELEGLAACEGEYSQKYSTMSPLGSGAFGFVW TAVDKEKNKEVVVKFIKKEKVLEDCWIEDPKLG KVTLEIAILSRVEHANIIKVLDIFENQGFFQLVME KHGSGLDLFAFIDRHPRLDEPLASYIFRQVRAG\Q SRLVSAVGYLRLKDIIHRDIKDENIVIAEDFTIKLI DFGSAAYLERGKLFYTFCGTIEYCAPEVLMGNPY

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RGPELEMWSLGVTLYTLVFEENPFCELEETVEAAIHPPYLVSKELMSLVSGLLQPVPERRTTLEKLVTDPWVTQPVNLADYTWEEVFRVNKPESGVLSAASLEMGNRSLSDVAQAQELCGGPVPGEAPNGQGCLHPGDPRLLTS
3862	A	399	2069	TMDRSKRNSIAGFPPRVE\RLEEFEGGGGGEGNVSQVGRVWPSSYRALISAFFRLTRLDDFTCEKIGSGFFSEVFKVRHRASGQVMALKMNTLSSNRANMLKEVQLMNRLSHPNILRYINSGNLEQLLDSNLHLPWTVRVKLAYDIAVGLSYLHFKGIFHRDLTSKNCLIKRDENGYSAVVADFGLAEKIPDVSMGSEKLAVVGSPFWMAPEVLRDEPYNEKADVFSYGIILCEIIARIQADPDYLPRTENFGLDYDAFQHMVGDCPPDFLQLTFNCCNMDPKLRPSFVEIGKTLEEILSRLQEEEQERDRKLQPTARGLLEKAPGVKRLSSLDDKIPHKSPCPRRTIWLSRSQSDIFSRKPPRTVSVLDPYYRPRDGAARTPKVNPFSARQDLMGGKIKFFDLPSKSVISLVFDLDAPGPGTMPLADWQEPLAPPIRRWRSLPGSPEFLHQEACPFVGREESLSDGPPPRLSSLKYRVKEIPPFRASALPAAQAHEAMDCSILQEENGFGSRPQGTSPCPAGASEEMEVEERPAGSTPATFSTSGIGLQTQGKQDG
3863	A	399	2069	TMDRSKRNSIAGFPPRVE\RLEEFEGGGGGEGNVSQVGRVWPSSYRALISAFFRLTRLDDFTCEKIGSGFFSEVFKVRHRASGQVMALKMNTLSSNRANMLKEVQLMNRLSHPNILRYINSGNLEQLLDSNLHLPWTVRVKLAYDIAVGLSYLHFKGIFHRDLTSKNCLIKRDENGYSAVVADFGLAEKIPDVSMGSEKLAVVGSPFWMAPEVLRDEPYNEKADVFSYGIILCEIIARIQADPDYLPRTENFGLDYDAFQHMVGDCPPDFLQLTFNCCNMDPKLRPSFVEIGKTLEEILSRLQEEEQERDRKLQPTARGLLEKAPGVKRLSSLDDKIPHKSPCPRRTIWLSRSQSDIFSRKPPRTVSVLDPYYRPRDGAARTPKVNPFSARQDLMGGKIKFFDLPSKSVISLVFDLDAPGPGTMPLADWQEPLAPPIRRWRSLPGSPEFLHQEACPFVGREESLSDGPPPRLSSLKYRVKEIPPFRASALPAAQAHEAMDCSILQEENGFGSRPQGTSPCPAGASEEMEVEERPAGSTPATFSTSGIGLQTQGKQDG
3864	A	3	911	SWNMDSDSCAAAFHPEEYSPSCKRRRTVEDFNKFCTFVLAYAGYIPYPKEELPLRSSPSPANSTAGTIDSDGWDAGFSDIASSVPLPVSDRCFSHLQPTLLQRAKPSNFLLDRKKTDKLKKKKKRKRRDSDAPGKEGYRGGLLKLEAADPYVETPTSPTLQDIPQAPSDPCSGWDSDTPSSGSCATVSPDQVKEIKTEGKRTIVR/QEAQLMARNDGNFSSLLESIFPS\DDDSWDLVTCFCMKPFAGRPMIECNECHTWIHLSCAKIRKSNVPEVFVCQKCRDSKFDIRRSNRSRTGSRKLFLD
3865	A	3	3573	QERLRSRSRPDRAAREAGSARGRQPKRTERVEQFLTIARRRGRRSMPVSLEDSGEPTSCPATDAETASEGSVESASETRSGPQSASTAVKERPASSEKVKGGDDHDDTSDSDSDGLTLKELQNRLRRKREQEPTERPLKGIQSRLRKKRREEGPAETVGSEASDTVEGVLPSKQEPENDQGVVSQAGKDDRESKLEGKAAQDIKDEEPGDLGRPKPECEGYDPNALYCICRQPHNN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RFMICCDRCEEWFHGDCVGISEARGRLLERNGE DYICPNCTILQVQDETHSETADQQEAKWRPGDA DGTDCTSIGTIEQKSSEDQGIKGRIEKAANPSGKK KLKIFQPGPGPVPTQLPVLWQVLEIAVSRSISAFT LLHCISCKVIEAPGASKCIGPGCCHVAQPDSVYCS NDCILKHAAATMKFLSSGKEQKPKPKEKMKMK PEKPSLPKCGAQAGIKISSVHKRPAPEKKETTVK KAVVVPARSEALGKEAACESSTPSWASDHNYNA VKPEKTAAPSPSLLYKSTKEDRRSEEKAAATAAS KKTAPPGSTVGKQPAPRNLVPKKSSFANVAAAT PAIKKPPSGFKGTIPKRPWLSATPSSGASAARQAG PAPAAATAASKKFPGSAALVGAVRKPVVPSVPM ASPAPGRLGAMSAAPSQPNSQIRQNIRRSLKEIL WK/RFLFFILFRVNDSDDLIMTENEVGKIALHIEK EMFNLFQVTDN/RAYKSKYRSIMFNLKDPKNQG LFHRVLREEISLAKLVRLKPEELVSKELSTWKER PARSVMESRTKLHNESKKTAPRQEAIPDLEDSPP VSDSEEQQESARAVPEKSTAPLLDVFSSMLKDTT SQHRAHLFDLNCKICTGQVPSAEDEPAPKKQKLS ASVKKEDLKSKHDSSAPDPAPDSADEVMPEAVP EVASEPGLESASHPNVDRTYFPGPPGDGHPEPSPL EDLSPCPASCGSGVVTTVTVSGRDPRTAPSSSCT AVASAASRPDSTHMVEARQDVPKPVLTSVMVPK SILAKPSSSPDPRYLSVPPSPNISTSESRSPPEGDTT LFLSRLSTIWKGFINMQSVAKFVTKAYPVSGCFD YLSEDLPDTIHIGGRIAPKTVWDYVGKLKSSVSK ELCLIRFHPATEEEEVAYISLYSYFSSRGRFGVVA NNNRHVKDLYLIPLSAQDPVPSKLLPFEGPGKRR LSGWR
3866	A	2	3181	AQQPVGRRGGASGAGGGRRGTPRPRAGAGPGF QVSSGGCRLSKMRRFLRPGHDPVRERLKRDLFQ FNKTVEHGFPHQPSALGYSPSLRILAIGTRSGAIK LYGAPGVEFMGLHQENNAVTQIHLLPGQCQLVT LLDDNSLHLWSLKVKGGASELQEDESFTLRGPP GAAPSATQITVVLPHSSCELLYLGTESGNVFVVQ LPAFRALEDRTISSDAVLQRLPEEARHRRVFEMV EALQEHPRDPNQILIGYSRGLVVIWDLQGSRVLY HFLSSQQLENIWWQRDGRLLVSCHSDGSYCQWP VSSEAQQPEPLRSLVPYGPFPCKAITRILWLTTRQ G\LPFTIFQGGMPRASYGDRHCISVIHDGQQTAFD FTSRVIGFTVLTEADPAATFDDPYALVVLAEEEL VVIDLQTAGWPPVQLPYLASLHCSAITCSHHVSN IPLKLWERIIAAGSRQNAHFSTMEWPIDGGTSLTP APPQRDLLLTGHEDGTVRFWDASGVCLRLLYKL STVRVFLTDTDPNENLSAQGEDEWPPLRKVGSF DPYSDDPRLGIQKIFLCKYSGYLAVAGTAGQVLV LELNDEAAEQAVEQVEADLLQDQEGYRWKGHE RLAARSGPVRFEPGFQPFVLVQCQPPAVVTSLAL HSEWRLVAFGTSHGFGLFDHQQRRQVFVKCTLH PSDQLALEGPLSRVKSLKKSLRQSFRRMRRSRVS SRKRHPAGPPGEAQEGSAKAERPGLQNMELAPV QRKIEARSAEDSFTGFVRTLYFADTYLKDSSRHC PSLWAGTNGGTIYAFSLRVPPAERRMDEPVRAE QAKEIQLMHRAPVVGILVLDGHSVPLPEPLEVAH DLSKSPDMQGSHQLLVVSEEQFKVFTLPKVSAK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LKLKLTALEGSRVRRVSVAHFGSRRAEDYGEHH LAVLTNLGDIQVVSLPLLKPQVRYSCIRREDVSGI ASCVFTKYGQGFYLISPSEFERFSLSTKG\LVEPRC LVDSAETKNHRPGNGAGPKKAPSRARNSGTQSD GEEKQPGLVMERALLSDERAATG\VHIEPPWGA ASAMAEQSEWLSVQAAR
3867	A	2	3181	AQQPVGRRGGASGAGGGRRGTPRPRAGAGPGF QVSSGGCRLSKMRRFLRPGHDPVRERLKRDLFQ FNKTVEHGFPHQPSALGYSPSLRILAIGTRSGAIK LYGAPGVEFMGLHQENNAVTQIHLLPGQCQLVT LLDDNSLHLWSLKVKGGASELQEDESFTLRGPP GAAPSATQITVVLPHSSCELLYLGTESGNVFVVQ LPAFRALEDRTISSDAVLQRLPEEARHRRVFEMV EALQEHPRDPNQILIGYSRGLVVIWDLQGSRVLY HFLSSQQLENIWWQRDGRLLVSCHSDGSYCQWP VSSEAQQPEPLRSLVPYGPFPCKAITRILWLTTRQ G\LPFTIFQGGMPRASYGDRHCISVIHDGQQTAFD FTSRVIGFTVLTEADPAATFDDPYALVVLAEEEL VVIDLQTAGWPPVQLPYLASLHCSAITCSHHVSN IPLKLWERIIAAGSRQNAHFSTMEWPIDGGTSLTP APPQRDLLLTGHEDGTVRFWDASGVCLRLLYKL STVRVFLTDTDPNENLSAQGEDEWPPLRKVGSF DPYSDDPRLGIQKIFLCKYSGYLAVAGTAGQVLV LELNDEAAEQAVEQVEADLLQDQEGYRWKGHE RLAARSGPVRFEPGFQPFVLVQCQPPAVVTSLAL HSEWRLVAFGTSHGFGLFDHQQRRQVFVKCTLH PSDQLALEGPLSRVKSLKKSLRQSFRRMRRSRVS SRKRHPAGPPGEAQEGSAKAERPGLQNMELAPV QRKIEARSAEDSFTGFVRTLYFADTYLKDSSRHC PSLWAGTNGGTIYAFSLRVPPAERRMDEPVRAE QAKEIQLMHRAPVVGILVLDGHSVPLPEPLEVAH DLSKSPDMQGSHQLLVVSEEQFKVFTLPKVSAK LKLKLTALEGSRVRRVSVAHFGSRRAEDYGEHH LAVLTNLGDIQVVSLPLLKPQVRYSCIRREDVSGI ASCVFTKYGQGFYLISPSEFERFSLSTKG\LVEPRC LVDSAETKNHRPGNGAGPKKAPSRARNSGTQSD GEEKQPGLVMERALLSDERAATG\VHIEPPWGA ASAMAEQSEWLSVQAAR
3868	A	1	2497	GDSGGPLVCEEPSGRFFLAGIVSWGIGCAEARRP GVYARVTRLRDWILEATTKASMPLAPTMAPAPA APSTAWPTSPESPVVSTPTKSMQALSTVPLDWVT VPKLQECGARPAMEKPTRVVGGFGAASGEVPW QVSLKEGSRHFCGATVVGDRWLLSAAHCFNHT KVEQVRAHLGTASLLGLGGSPVKIGLRRVVLHP LYNPGILDFDLAVLELASPLAFNKYIQPVCLPLAI QKFPVGRKCMISGWGNTQEGNATKPELLQKASV GIIDQKTCSVLYNFSLTDRMICAGFLEGKVDSCQ VSGIKALYESELADARRVLDETARERARLQIEIG KLRAELDEVNKSAKKREGELTVAQGRVKDLESL FHRSEVELAAALSDKRGLESDVAELRAQLAKAE DGHAVAKKQLEKETLMRVDLENRCQSLQEELDF RKSVFEEEVRETRRRHERRLVEVDSSRQQEYDFK MAQALEELRSQHDEQVRLYKLELEQTYQAKLDS AKLSSDQNDKAASAAREELKEARMRLESLSYQL SGLQKQASAAEDRIRELEEAMAGERDKFRKMLD

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AKEQEMTEMRDVMQQQLAEYQELLDVKLALD MEINAYRKLLEGEEERLKLSPSPSSRVTVSRATSS SSGSLSATGRLGRSKRKR\WRWRSPW\QRPKRPG HGHGWQRWLPPGPAGLGLGQR\HIEEIDLEGKFV QLKNNSDKDQSLGNWRIKRQVLEGEEIAYKFTP KYILRAGQMVTVWAAGAGVAHSPPSTLVWKGQ SSWGTGESFRTVLVNADGEEVAMRTVKKSSVM RENENGEEEEEEAEFGEEDLFHQQGDPRTTSRGC YVM
3869	A	1	1942	RYRAGIPGDGRKDYIRLTRPGLTLPGRAMFARGS RRRRSGRAPPEAEDPDRGQPCNSCREQCPGFLLH GWRKICQHCKCPREEHAVHAVPVDLERIMCRLIS DFQRHSISDDDSGCASEEYAWVPPGLKPEQVYQ FFSCLPEDKVPYVNSPGEKYRIKQLLHQLPPHDS EAQYCTAL\EE\EEKKELRAFSQQRKRENLG/RLG IVRIFPVTIT\GAI\CEECGKQIGGGDIAVF\ASRASL GLLLGQPSCF\VCTTCQELLVDLIYFYHVGKVYC GRHHAECLRPRCQACDEIIFSPECTEAEGRHWHM DHFCCFECEASLGGQRYVMRQSRPHCCACYEAR HAEYCDGCGEHIGLDQGQMAYEGQHWHASDRC FCCSRCGRALLGRPFLPRRGLIFCSRACSLGSEPT APGPSRRSWSAGPVTAPLAASTASFSAVKGASET TTKGTSTELAPATGPEEPSRFLRGAPHRHSMPEL GLRSVPEPPPESPGQPNLRPDDSAFGRQSTPRVSF RDPLVSEGGPRRTLSAPPAQRRRPRSPPPRAPSRR RHHHHNHHHHHNRHPSRRRHYQCDAGSGSDSE SCSSSPSSSSSESSEDDGFFLGERIPLPPHLCRPMP AQDTAMETFNSPSLSLPRDSRAGMPRQARDKNC IVA
3870	A	2	3485	FVWRVFYVHASCMPPRARSWEGAHAPVGMHV AEAHACSSQQQQMPPAQFWMLEWLLHLCAFLS TPSFPHWCCCSNPHGSIADKPEEIVPASKPSRAAE NMAVEPRVATIKQRPSSRCFPAGSDMNSVYERQ GIAVMTPTVPGSPKAPFLGIPRGTMRRQKSIDSRI FLSGITEEERQFLAPPMLKFTRSLSMPDTSEDIPPP PQSVPPSPPPPSPTTYNCPKSPTPRVYGTIKPAFNQ NSAAKVSPATRSDTVATMMREKGMYFRRELDR YSLDSEDLYSRNAGPQANFRNKRGQMPENPYSE VGKIASKAVYYVPAKPARRKGMLVKQSNVEDSPE KTCSIPIPTIIVKEPSTSSSGKSSQGSSMEIDPQAPE PPSQLRPDESLTVSSPFAAAIAGAVRDREKRLEA RRNSPAFLSADLGDEHVGLGPPAPRTRPSMFPEE GDFADEDSAEQLSSPMPSATPREPENHFVGGAEA SAPGEAGRPLNSTSKAQGPESSPAVPSASSGTAG PGNYVHPLTGRLLDPSSPLALALSARDRAMKES QQGPKGEAPKADLNKPLYIDTKMRPSLDAGFPT VTRQNTRGPLRRQETENKYETDLGRDRKGDDK KNMLIDIMDTSQQKSAGLLMVHTVDATKLDNA LQEEDEKAEVEMKPDSSPSEVPEGVSETEGALQI SAAPEPTTVPGRTIVAVGSMEEAVILPFRIPPPPLA SVDLDEDFIFTEPLPPPLEFANSFDIPDDRAASVPA LSDLVKQKKSDTPQSPSLNSSQPTNSADSKKPAS LSNCLPASFLPPPESFDAVADSGIEEVDSRSSSDH HLETTSTISTVSSISTLSSEGGENVDTCTVYADGQ AFMVDKPPVPPKPKMKPIIHKSNALYQDALVEE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DVDSFVIPPPAPPPPPGSAQPGMAKVLQPRTSKLWGDVTEIKSPILSGPKANVISELNSILQQMNREKLAKPGEGLDSPMGAKSASLAPRSPEIMSTISGTRSTTVTFTVRPGTSQPITLQSRPPDYESRTSGTRRAPSPVVSPTEMNKETLPAPLSAATASPSPALSDVFSLPSQPPSGDLFGLNPAGRSRSPSPSILQQPISNKPFTTKPVHLWTKPDVADWLESLNLGEHKEAFMDNEIDGSHLPNLQKEDLIDLGVTRVGHRMNIERALKQLLDR
3871	A	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDVKKKIKEVTEEVANKVSCAMTDEICRLSVLVDEFCSEFHPNPDVLKIYKSELNKHIEDGMGRNLADRCTDEVNALVLQTQQEIIENLKPLLPAGIQDKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIVFRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIFQLPRSLASTPTAPTTPATPDNASQEELMITLVTGLASVTSRTSMGIIIVGGVIWKTIGWKLLSVSLTMYGALYLYERLSWTTHAKERAFKQQFVNYATEKLRMIVSSTSANCSHQVKQQIATTFARLCQQVDITQKQLEEEIARLPKEIDQLEKIQNNSKLLRNKAVQLENELENFTKQFLPSSNEES
3872	A	35	1171	VESRSAWHEGEDQIDRLDFIRNQMNLLTLDVKKKIKEVTEEVANKVSCAMTDEICRLSVLVDEFCSEFHPNPDVLKIYKSELNKHIEDGMGRNLADRCTDEVNALVLQTQQEIIENLKPLLPAGIQDKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIVFRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIFQLPRSLASTPTAPTTPATPDNASQEELMITLVTGLASVTSRTSMGIIIVGGVIWKTIGWKLLSVSLTMYGALYLYERLSWTTHAKERAFKQQFVNYATEKLRMIVSSTSANCSHQVKQQIATTFARLCQQVDITQKQLEEEIARLPKEIDQLEKIQNNSKLLRNKAVQLENELENFTKQFLPSSNEES
3873	A	2944	2089	PVCTALTPGRMTDDKDVLRDVWFGRIPTCFTLYQDEITEREAEPYYLLLPRVSYLTLVTDKVKKHFQKVMRQEDISEIWFEYEGTPLKWHYPIGLLFDLLASSSALPWNITVHFKSFPEKDLLHCPSKDAIEAHFMSCMKEADALKHKSQVINEMQKKDHKQLWMGLQNDRFDQFWAINRKLMEYPAEENGFRYIPFRIYQTTTERPFIQKLFRPVAADGQLHTLGDLLKEVCPSAIDPEDGEKKNQVMIHGIEPMLETPLQWLSEHLSYPDNFLHISIIPQPTD
3874	A	776	366	QARGAPSSPMCPLPLAAAAVAAPRAPLRLLNRGLAAAMSTAQSLKSVDYEVFGRVQGVCFRMYTEDEARKIGVVGWVKNTSKGTVTGQVQGPEDKVNSMKSWLSKVGSPSSRIDRTNFSNEKTISKLEYSNFSIRY
3875	A	1081	182	SLSSCQTDPRPMSAPLDAALHALQEEQARLKMRLWDLQQLRKELGDSPKDKVPFSVPKIPLVFRGHTQQDPEVPKSLVSNLRIHCPLLAGSALITFDDPKVAEQVLQQKEHTINMEECRLRVQVQPLELPMVTTIQVMVSSQLSGRRVLVTGFPASLRLSEEELLDKLEIFFGKTRNGGGDVDVRELLPGSVMLGFARDGVAQRLCQIGQFTVPLGGQQVPLRVSPYVNGEIQKAEIRSQPVPRSVLVLNIPDILDGPELHDVLEIHFQKPT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				RGGGEVEALTVVPQGQQGLAVFTSESG
3876	A	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPLSLVVHEGDTVTLNCSYEVTNFRSLLWYKQEKKAPTFLFMLTSSGIEKKSGRLSSILDKKELSSILNITATQTGDSAIYLCAVEAQCSLVTCSLYSNSTAEALQL
3877	A	3	1291	KAFRLLAERGAAAAMLWSGCRRFGARLGCLPGGLRVLVQTGHRSLTSCIDPSMGLNEEQKEFQKVAFDFAAREMAPNMAEWDQKELFPVDVMRKAAQLGFGGVYIQTDVGGSGLSRLDTSVIFEALATGCTSTTAYISIHNMCAWMIDSFGNEEQRHKFCPPLCTMEKFASYCLTEPGSGSDAASLLTSAKKQGDHYILNGSKAFISGAGESDIYVVMCRTGGPGPKGISCIVVEKGTPGLSFGKKEKKVGWNSQPTRAVIFEDCAVPVANRIGSEGQGFLIAVRGLNGGRINIASCSLGAAHASVILTRDHLNVRKQFGEPLASNQYLQFTLADMATRLVAARLMVRNAAVALQEERKDAVALCSMAKLFATDECFAICNQALQMHGGYGYLKDYAVQQYVRDSRVHQILEGSNEVMRILISRSLLQE
3878	A	10	1014	LPGSTISSSGCQAPGRADSSGGARNSRRGDSRPGSCNRQAVAPPCPSPGPQSRHWIHRGTAPQAGETRTLGRGSSAPNACSASVTPCCPSSPPS*SCL*PTRRSPQNSSSTEVYRGFWQHGLPST**PFSS*QWPGQHTQGCSKLLGKQTTHLPCSTWPA**PSPSCLTRFR*W*PSLMCLWASSCSVCV*SPSGSCRH*LWGTHSTSRTC*ARRSSALPTGLCTDDTSWASSSKARPCALQRPSSLSSLSPCLTC*W*LSSSSPMSARSPAGAETGSWATGSPRLTQWKSSRLTSTSHSARSAWKPSATESTPSWPRFSSWTSGEDPASPAPAI
3879	A	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSAPGNTSLCTRDYKITQVLFPLLYTVLFFVGLITNGLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTFPFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYISISFLGLITIDRYQKTTRPFKTSNPKNLLGAKILK
3880	A	26	169	QPETDTMVHLTPEEKSAVTALWGKVNVDEDAGDDLCQILVDRPRLRI
3881	A	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPPACRIMPTTVDDVLEHGGEFHFFQKQMFFLLALLSATFAPIYVGIVFLGFTPDHRCRSPGVAELSLRCGWSPAEELNYTVPGPGPAGEASPRQCRRYEVDWNQSTFDCVDPLASLDTNRSRLPLGPCRDGWVYETPGSSIVTEFNLVCANSWMLDLFQSSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLINAAAGVLMAISPTYTWMLIFRLIQGLVSKAGWLIGYILITEFVGRRYRRTVGIFYQVAYTVGLLVLAGVAYALPHWRWLQFTVALPNFFFLLYYWCIPESPRWLISQNKNAEAMRIIKHIAKKNGKSLPASL
3882	A	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAEAMLDEPQEQAEGSLTVYVISEHSSLLPQDMMSYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQAMSLTEDVLAAALADHLPEDKWSAEKRRPLKSSLGYEITFSLLNPDPKSHDVYWDIEGAVRRYVQPFLNALGAAGNFSVDSQILYYAMLGVNPRFDSASSSYYLDMHSLPHVINPVESRLGSSAASLYPVLNFLLYVPELAHSPLYIQDKDGAPVATNAFHSPRWGGIMVYNVDSKTYNASVLPVRVEVDMVRVMEVFLAQLRLLFGI

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				AQPQLPPKCLLSGPTSEGLMTWELDRLLWARSV ENLATATTTLTSLA
3883	A	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTKKIQ FHQELLVLFWKLCDFNKVGQPRGALQGDGEQLP Q*PGGRDSVRLRGVGQSCPSLELSPLGPSPHP*KF LFFVLKSSDVLDILVPILFFLNDARADQSRVGLM HIGVFILLLLSGECNFGVRLNKPYSIRVPMDIPVF TGTHADLLIV\VFHKIITSGHQRLQPLFDCLLTIVV NVSPYLKSLSMVTANKLLHLLEAFSTTWFLFSAA QNHHLVFFLLEVFNNIIQYQFDGNSNLVYAIIRKR SIFHQLANLPTDPPTIHKALQRRRRTPEPLSRTGS QGGAPPWRAPAPLPLQSQAPSRPVWWLLQALTS *PRSPRCQRMAPCGPWNLSPSRAWRMAARLRGS PARHGGSSGDRP/HSSASGQWSPTPEWVLSWKS KLPLQTIMRLLQVLVPQVEKICIDKGLTDESEILR FLQHGTLVGLLPVPHPILIRKYQANSGTAMWFRT YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
3884	A	1	804	NGPRAPFSQEGQSTGPPPLIPRLGQHGAQGRIPPL NPGQGPGPNKDDSRGPPNHHMGPMSERRHEQSG GPEHGPERGPLRGGQDCRGPPDRRGPHPDFPDDF SRPDDFHPDKRFGHRLREFEGRGGPLPQEEKWR RGGPGPPFPPDHREFSEGDGRGAARGPPGAWEG RRPGG*TFPPGSRGPTFS/SGAEEESFRRGAPPRHE GRAPPRGRDGFPGPEDFGPEENFDASEEAARGRD LRGRGRGTPRGERVTKDTWSGRIGCRIHWL
3885	A	3	996	GRRRAGPAHSARMYNMMETELKPPGPQQTSGG GGGNSTAAAAGGNQKNSPDRVKRPMNAFMVW SRGQRRKMAQENPKMHNSEISKRLGAEWKLLSE TEKRPFIDEAKRLRALHMKEHPDYKYRPRRKTK TLMKKDKYTLPGGLLAPGGNSMASGVGVGAGL GAGVNQRMDSYAHMNGWSNGSYSMMQDQLG YPQHPGLNAHGAAQMQPMHRYDVSALQYNSM TSSQTYMNG/SRPTYSMSYSQQGTPGMAPGS\MG SVVKSEASSSPPVVTSSSHSRAPCQAGDLRDMIS MYLPGAEVPEPAAPSRLHMSQHYQSGPVPGTAI NGTLPLSHM
3886	A	773	317	QCTQKAAEGYTQFYYVDVLDGKLACVNKCTKG TKSQMNCNLGTCQLQRSGPRCLCPNTNTHWYW GETCEFNIAKSLVYGIVGAVMAVLLLALIILIILFS LSQ\RKRHRPESEGEADFGLENATNNFG\PTLETV DSGTELHIQ\RPEMVASTV
3887	A	3	466	VDFRVKTLLVDNKCFVLQLWDTAGQERYHSMT RQLLRKADGVVLMYDITSQESFAHVRYWLDCL QDAGSDGVVILLLGNKMDCEEERQVSVEAGQQL AQELGVYFGECSAALGHNILEPVVNLARSLRMQ EEGLKDSLVKVAPKRPPKRFGCCS
3888	A	3412	3144	QNIDITNFSSSWNDGLAFCALLHTYLPAHIPYQEL NSQDKRRNFMLAFQAAESVGIKSTLDINEMVRT ERPDWQNVMLYVTAIYKYFET
3889	A	1	1160	LVVTAITAILAFPNEYTRMSTSELISELFNDCGLL DSSKLCDYENRFNTSKGGELPDRPAGVGVYSAM WQLALTLILKIVITIFTFGMKIPSGLFIPSMAVGAI AGRLLGVGMEQLAYYHQEWTVFNSWCSQGAD CITPGLYAMVGAAACLGGVTRMTVSLVVIMFEL TGGLEYIVPLMAAAMTSKWVADALGREGIYDA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				HIRLNGYPFLEAKEEFAHKTLAMDVMKPRRNDP LLTVLTQDSMTVEDVETIISETTYSGFPVVVSRES QRLVGFVLRRDLIISIENARKKQDGVVSTSIIYFTE HSPPLPPYTPPTLKLRNILDLSPFTVTDLTPMEIVV DIFRKLGLRQCLVTHNGRLLGIITKKDVLKHIAQ MANQDPDSILFN
3890	A	1	387	SWCWTGIFVLGTTNLRLEGSWYRSLWGPGFNTT TATLGFGAPQAPVGDVALNQPDMCVYRRGRKK RVPYTKLQLKELENEYAINKFINKDKRRRISAAT NLSERQVTIWFQNRRVKDKKIVSKLKDTVS
3891	A	2	2914	RGGGGDHKMADLSLLQEDLQEDADGFGVDDYS SESDVIIIPSALDLAST/QDEMVERPLGRL\DK\YA ASENHI*PDKMVAPEFASIPLRE\VCDDERDCIAV LGKN*PDWADDSEPT\VRAAELEQVPHIALFLFK KTRLSITICFFSKFLLPYCGLDTLADQN\NQVRKT SQAALL\ALLEQELIERFDVETKVCPVLIELTAPDS NDDVKTEAVAIMCKMAP\MVGKDITERLILPRFC EMCCDCRMFH\VRK\VCAANFGDICSVVGQQAT EEMLLPRFFQLCSDNVWGVRKACAECFMAVSC ATCQEIRRTKLSALFINLISDPSRWVRQAAFQSLG PFISTFANPSSSGQYFKEESKSSEEMSVENNKRTR DQEAPEDVQVRPEDTPSDLSVSNSSVILENTMED HAAEASGKPLGEISVPLDSSLLCTLSSESHQEAAS NENDKKPGNYKSMLRPEVGTTSQDSALLDQELY NSFHFWRTPLPEIDLDIELEQNSGGKPSPEGPEEE SEGPVPSSPNITMATRKELEEMIENLEPHIDDPDV KAQVEVLSAALRASSLDAHEETISIEKRSDLQDE LDINELPNCKINQEDSVPLISDAVENMDSTLHYIH NDSDLSNNSSFSPDEERRTKVQDVVPQALLDQY LSMTDPSRAQTVDTEIAKHCAYSLPGVALTLGR QNWHCLRETYETLASDMQWKVRRTLAFSIHELA VILGD\QLTAADLVPIFNGFLK*PSMKSRIGVLKH LHDFLKLLHIDKRREYLYQLQEFLVTDNSRNWR FRAELAEQLILLLELYSPRDVYDYLRPIALNLCAD KVSSVRWISYKLVSEMVKKLHAATPPTFGVDLIN ELVENFGRCPKWSGRQAFVFVCQTVIEDDCLPM DQFAVHLMPHLLTLANDRVPNVRVLLAKTLRQT LLEKDYFLASASCHQEAVEQTIMALQMDRDSDV KYFASIHPASTKISEDAMSTASSTY
3892	A	158	2191	VPLPAPSGLSGGGSRGAGCKKAPPGRAPAPGLAP LRPSEPTMAVPPGHGPFSGFPGPQEHTQVLPDVR LLPRRLPLAFRDATSAPLRKLSVDLIKTYKHINEV YYAKKKRRAQQAPPQDSSNKKEKKVLNHGYDD DNHDYIVRSGERWLERYEIDSLIGKGSFGQVVKA YDHQTQELVAIKIIKNKKAFLNQAQIELRLLELM NQHDTEMKYYIVHLKRHFMFRN\HLCLVFELLS YNLYDLLRNTHFRGVSLNLTRKLAQQLCTALLF LATPELSIIHCDLKPENILLCNPKRSAIKIVDFGSS CQLGQRIYQYIQSRFYRSPEVLLGTPYDLAIDMW SLGCILVEMHTGEPLFSGSNEVCPQEGVDQMNRI VEVLGIPPAAMLDQAPKARKYFERLPGGGWTLR RTKELRKDYQGPGTRRLQEVLGVQTGGPGGRRA GEPGHSPAD\Y\LRFQDLVLRMLEYEPAARISPLG ALQHGFFRRTADEATNTGPAGSSASTSPAPLDTC PSSSTASSISSSGGSSGSSSDNRTYRYSNRYCGGP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				GPPITDCEMNSPQVPPSQPLRPWAGGDVPHKTH QAPASASSLPGTGAQLPPQPRYLGRPPSPTSPPPP ELMDVSLVGGPADCSPPHPAPAPQHPAASALRT RMTGGRPPLPPPDDPATLGPHLGLRGVPQSTAAS S
3893	A	68	258	PEEYYPFSPTLQQLFFFLLDSDMGSRPESMGCRK NTVPRPASPTEAGTDPQTFLHTWVSECRD
3894	A	1120	136	SLPLAPAPAVAGPVALCPAGLCPAQPGMPAGPA AASGSHPEVGSVLQRSSQPHWPNPWPGAGHLPP PAGPFPYNPPAGPGAAAGLA*SPPRSSPTPCSVGP QSCPANASAPPAQPCLAGAPPAASLPPPGPGSVS AAPAPGGPAPAEPPLGVPPVPAWLLPDSPPLPGT HSGPPPAAVSLPPAAAACPVVVPPPLPHHPPDLES PSAAAPNPGCAGGIRHFPPGSPEASSPLRPAAAPA LLPLPRPPS*P/VPWKPLHSPVAVAGGSFVAGGSV LPAPDLDQPRPSGPPAASPTPGPGVAQPPPGSAVL PTVP*APPVSGAAPGRKREW
3895	A	2	1347	FGAVSYRPGNGSCWVKVTASSDLSDLISCLCPPR SLCSSQACVLPVPGPSLLLPQGLHVGCASAGTRW PLSCSIDFQRLLAHEEETQKRRAKESGMAFTQLT FRDVAIEFSQDEWKCLNSTQRTLYRDVMLENYR NLVSLDLSRNCVIKELAPQQEGNP/ARSIPHSDIGT T*KT*H*RVLLQGNQEKNTRL*LSVER**KKLQQ SDYGPKRKSYL*ERPTR*KRYRKQVY*TSA*LSF LPHPHELQQFQAEGKIYECNHVEKSVNHGSSVSP PQIISSTIKTHVSNKYGTDFICSSLLTQEQKSCIRE KPYRYIECDKALNHGSHMTVRQVSHSGEKGYKC DLCGKVFSQKSNLARHWRVHTGEKPYKCNECD RSFSRNSCLALHRRVHTGEKPYKCYECDKVFSR NSCLALHQKTHIGEKPYTCKECGQAFSVRSTLTN HQVIHSDK
3896	A	202	498	MVQSCSAYGCKNRYDKDKPVSFHKFPLTRPSLC KEWEAAVRRKNFKPTKYSSICSEHFTPDCFKREC NNKLLKENAVPTIFLCTEPHDKKEDLLEPQEQ
3897	A	2	382	SHGLSRAPHLSAAPAPALASRPCFSSAPCSQGGG GGGPATMIHFILLFSRQGKLRLQKWYITLPDKER KKITREIVQIILSRGHRTSSFVDWKELKLVYKRYA SLYFCCAIE\NQDNELLTLENVHR
3898	A	718	305	SEQEPLLGDTPGSREWDILETEEHYKSRWRSIRIL YLTMFLSSVGFSVVMMSIWPYLQKIDPTADTSFL GWVIASYSLGQMVASPIFGLWSNYRPRKEPLIVSI LISVAANCLYAYLHIPASHNKYYMLVARGLLGIG
3899	A	24	718	FRGRPGIPEREGKGNHSFVEVARVIVVDLHSRLG GAMAERKGTAKVDFLKKIEKEIQQKWDTERVFE VNASNLEKQTSKGKYFVTFPYPYMNGRLHLGHT FSLSKCEFAVGYQRLKGKCCLFPFGLHCTGMPIK ACADKLKREIELY/GCPPDFPDEEEEEEETSVKTE DIIKDKAKGKKSKAA/AKAGSSKYQWGIMKSLG LSDEEIVKFSEAEHWLDYFNALAIQDLKRMG
3900	A	360	1	VPATSSNVSPSSSESSEPDLSSRSSSSDAPSSSPSVP SPCSLSLSSPESPLLPTLLSSKSPAGSAGPTCGCPS GPGLRATA/PSRLSSSIAAH/SSSAPETSRPAAARE RSPPLHDRESHE
3901	A	193	345	GEWAVPPAPGGQGVSIPHGPEPGQGSGVHIAPRQ GEGSDRTEPLICPKAAP

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3902	A	1188	1389	NPAARSAAAREGSPALPPPPVS/SSSGLGLLLPLSPPGSHAANPALSPRAPHSHYRPRPRCGPRRRPR
3903	A	63	396	NNMRNPHLSSNHYLNLARTETVFARMESVKQRILAPGKEGLKNFAGKSLGQIYRVLEKKQDTGETIELTEDGKPL*VPERKAPLCDCTCFGLPRRYIIAIMSGLGFCISFG
3904	A	732	1046	AMSECPLILYIHKHIDTYSQSYLFNDLFYPVYSGGRMVTYEHLREVVFGKSEDEHYPLW*VLFGK*YAVAPNALMFIRFM*NCTFVPKLP*VMDLK**LQYKSR
3905	A	46	910	QPPPPPPPPPSPPPPPFPPARALSHLRLHPDACLFPSPFPLPCSTMPGMMEKGPELLGKNRSANGSAKSPAGGGGSGASSTNGGLHYSEPESGCSSDDEHDVGMRVGAEYQARIPEFDPGATKYTDKDNGGMLVWSPYHSIPDAKLDEYIAIAKEKHGYNVEQALGMLFWHKHNIEKSLADLPNFTPFPDEWTVEDKVLFEQAFSFHGKSFHRIQQMLPDKTIASLVKYYYSWKKTRSRTSLMDRQARKLANRHNQGDSDDDVEETHPMDGNDSDYDPKKEAKKEGMS
3906	A	2	513	KVCNCCSQELETSFTYVDKNINLEQRNRSSPSAKGHNHPGELGWENPNEWSQEAAISLISEEEDDTSSEATSSGKSIDYGFISAILFLVTGILLVIISYIVPREVTVDPNTVAAREMERLEKESARLGAHLDRCVIAGLCLLTLGGVILSCLLMMSMWKGELYRRNRFAS
3907	A	71	412	ILIMSNCLQNFLKITSTRLLCSRLCQQLRSKRKFFGTVPISRLHRRVVITGIGLVTPLGVGTHLVWDRLIGGESGIVSLVGEEYKSIPCSVAAYVPRGSDEGQFNEQNFVSKSD
3908	A	77	746	LGTLLGWRAPLFSRCLAFHSPFILLNTPKLVKTAELPPDRNYVLGAHPHGIMCTGFLCNFSTESNGFSQLFPGLRPWLAVLAGLFYLPVYRDYIMSFGLCPVSRQSLDFILSQPQLGQAVVIMVGGAHEALYSVPGEHCLTLQKRKGFVRLALRHGASLVPVYSFGENDIFRLKAFATGSWQHWCQLTFKKLMGFSPCIFWGRGLFSATSWGLLPFAVPITTV
3909	A	1	793	FRAAGRPAAAMGDIPVVGLSSWKASPGKVTEAVKEAIDAGYRHFDCAYFYHNEREVGAGIRCKIKEGAVRREDLLIATKLWCTCHKKSLVETACRKSLKALKLNYLDLYLIHWPMGFKPPHPEWIMSCSELSFCLSHPRVQDLPLDESNMVIPSDTDFLDTWEAMEDLVITGLVKNIGVSNFNHEQLERLLNKPGLRFKPLTNQIECHPYLTQKNLISFCQSRDVSVTAYRPLGGSCEGVDLIDNPVIKRIAKEHGKSPAQILI
3910	A	202	705	FFTMHRKKVDNRIRILIENGVAERQRSLFVVVGDRGKDQVVILHHMLSKATVKARPSVLWCYKKELGFSSHRKKRMRQLQKKIKNGTLNIKQDDPFELFIAATNIRYCYYNETHKILGNTFGMCVLQDFEALTPNLLARTVETVEGGGLVVILLRTMNSLKQLYTVTM
3911	A	3	723	AGRGARAAGEGGGPFKSRPRPLPSSRSLPAVGGGRYGADKMAAGGAVAAAPECRLLPYALHKWSSFSSTYLPENILVDKPNDQSSRWSSESNYPPQYLILKLERPAIVQNITFGKYEKTHVCNLKKFKVFGGMNEENMTELLSSGLKNDYNKETFTLKHKIDEQMFPCRFIKIVPLLSWGPSFNFSIWYVELSGIDDPDIVQPC

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				LNWYSKYREQEAIRLCLKHFRQHNYTEAFESLQKKT
3912	A	2	461	FEKKQLRRPSLFLLGCCSFGIMAPSLWKGLEGIGLFALAHAAFSAAQHRSYMRLTEKEDESLPIDIVLQTLLAFAVTCYGIVHIAGEFKDMDATSELKNKTFDTVRNHPSFYVFNHRGSEYFSGPSDTANSSNQDALSSNTSLKLRKLESLRR
3913	A	362	20	APGRPEAKVPERSRESGSRRVRGPLLQLRPGRTSRPASGRGRGGAGGSYGKMRKPDSKIVLLGDMNVGKTSLLQRYMERRFPDTVSTVGGAFYLKQWRSYNISIWDTAGEAGAA
3914	A	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTEKQLQCMPMEGRGRASSSISDLQGKGFEKGTGEKHVPGVGSARHSPQASAGGSPWQRGKAQTRWLGKPDPGRKRRRGSPQEEGGLRVSAAARLLCSGANRCKVLVRQNSTPNTQQPAVHPSTPPSRPLPQAGRCLVAPLRPHPDWVAAKTLAKALRAPGKPWRLAAPSPLGDLGAPGLPGPSTAPRTLSVEEPGVECNQLCLYADVTDPVLCLGQKDPGVEGKHCEKEKISSSKELKHVHAKSEPSKPARRLSESLHVVDENKNESKIEREHKRRTSTPVIMEGVQEETDTRDVKRQVERSEICTEEPQKQKSTLKNEKHLKKDDSETPHLKSLLKKEVKSSKEKPEREKTPSEDKLSVKHKYKGDCMHKTGDETELHSSEKGLKVEENIQKQSQQTKLSSDDKTERKSKHRNERKLSVLGKDGKPVSEYIIKTDENVRKENNKKERRLSAEKTKAEHKSRRSSDSKIQKDSLGSKQHGITLQRRSESYSEDKCDMDSTNMDSNLKPEEVVHKEKRRTKSLLEEKLVLKSKSKTQGKQVKVVETELQEGATKQATTPKPDKEKNTEENDSEKQRKSKVEDKPFEETGVEPVLETASSSAHSTQKDSSHRAKLPLAKEKYKSDKDSTSTRLERKLSDGHKSRSLKHSSKDIKKKDENKSDDKDGKEVDSSHEKARGNSSLMEKKLSRRLCENRRGSLSQEMAKGEEKLAANTLSTPSGSSLQRPKKSGDMTLIPEQEPMEIDSEPGVENVFEVSKTQDNRNNNSHQDIDSENMKQKTSATVQKDELRTCTADSKATAPAYKPGRGTGVNSNSEKHADHRSTLTKKMHIQSAVSKMNPGEKEPIHRGTTEVNIDSETVHRMLLSAPSENDRVQKNLKNTAAEEHVAQGDATLEHSTNLDSSPSLSSVTVVPLRESYDPDVIPLFDKRTVLEGSTASTSPADHSALPNQSLTVRESEVLKTSDSKEGGEGFTVDTPAKASITSKRHIPEAHQATLLDGKQGKVIMPLGSKLTGVIVENENITKEGGLVDMAKKENDLNAEPNLKQTIKATVENGKKDGIAVDHVVGLNTEKYAETVKLKHKRSPGKVKDISIDVERRNENSEVDTSAGSGSAPSVLHQRNGQTEDVATGPRRAEKTSVATSTEGKDKDVTLSPVKAGPATTTSSETRQSEVALPCTSIEADEGLIIGTHSRNNPLHVGAEASECTVFAAAEEGGAVVTEGFAESETFLTSTKEGESGECAVAESEDRAADLLAVHAVKIEANVNSVVTEEKDDAVTSAGSEEKCDGSLSRDSEIVEGTITFISEVESDGAVTSAGTEIRAGSISSEEVDGSQGNMMRMGPKKETEGTVTCTGAEGRSDNFVICSVTGAGPREERMVTGAGVVLGDNDAPPGTSASQEGDGSVNDGTEGESAVTSTGITEDGEGPASCTGSEDSSEGFAISSESEEENGESA

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MDSTVAKEGTNVPLVAAGPCDDEGIVTSTGAKE EDEEGEDVVTSTGRGNEIGHASTCTGLGEESEGV LICESAEGDSQIGTVVEHVEAEAGAAIMNANENN VDSMSGTEKGSKDTDICSSAKGIVESSVTSAVSG KDEVTPVPGGCEGPMTSAASDQSDSQLEKVEDT TISTGLVGGSYDVLVSGEVPECEVAHTSPSEKED EDIITSVENEECDGLMATTASGDITNQNSLAGGK NQGKVLIISTSTTNDYTPQVSAITDVEGGLSDALR TEENMEGTRVTTEEFEAPMPSAVSGDDSQLTASR SEEKDECAMISTSIGEEFELPISSATTIKCAESLQP VAAAVEERATGPVLISTADFEGPMPSAPPEAESP LASTSKEEKDECALISTSIAEECEASVSGVVVESE NERAGTVMEEKDGSGIISTSSVEDCEGPVSSAVP QEEGDPSVTPAEEMGDTAMISTSTSEGCEAVMIG AVLQDEDRLTITRVEDLSDAAIISTSTAECMPISA SIDRHEENQLTADNPEGNGDLSATEVSKHKVPM PSLIAENNCRCPGPVRGGKEPGPVLAVSTEEGHN GPSVHKPSAGQGHPSAVCAEKEEKHGKECPEIGP FAGRGQKESTLHLINAEEKNVLLNSLQKEDKSPE TGTAGGSSTASYSAGRGLEGNANSPAHLRGPEQ TSGQTAKDSSVSSIRYLAAVNTGAIKADDMPPVQ GTVAEHSFLPAEQQGSEDNLKTSTTKCITGQESKI APSHTMIPPATYSVALLAPKCEQDLTIKNDYSGK WTDQASAEKTGDDNSTRKSFPEEGDIMVTVSSE ENVCDIGNEESPLNVLGGLKLKANLKMEAYVPS EEEKNGEILAPPESLCGGKPSGIAELQREPLLVNE SLNVENSGFRTNEEIHSESYNKGEISSGRKDNAE AISGHSVEADPKEVEEEERHMPKRKRKQHYLSSE DEPDDNPDVLDSRIETAQRQCPETEPHATKEENS RDLEELPKTSSETNSTTSRVMEEKDEYSSSETTGE KPEQNDDDTIKSQE
3915	A	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTEKQ LQCMPMEGRGRASSSISDLQGKGFEKGTGEKHV PGVGSARHSPQASAGGSPWQRGKAQTRWLGKP DPGRKRRRGSPQEEGGLRVSAAARLLCSGANRC KVLVRQNSTPNTQQPAVHPSTPPSRPLPQAGRCL VAPLRPHPDWVAAKTLAKALRAPGKPWRLAAP SPLGDLGAPGLPGPSTAPRTLSVEEPGVECNQLC LYADVTDPVLCLGQKDPGVEGKHCEKEKISSSK ELKHVHAKSEPSKPARRLSESLHVVDENKNESKI EREHKRRTSTPVIMEGVQEETDTRDVKRQVERSE ICTEEPQKQKSTLKNEKHLKKDDSETPHLKSLLK KEVKSSKEKPEREKTPSEDKLSVKHKYKGDCMH KTGDETELHSSEKGLKVEENIQKQSQQTKLSSDD KTERKSKHRNERKLSVLGKDGKPVSEYIIKTDEN VRKENNKKERRLSAEKTKAEHKSRRSSDSKIQK DSLGSKQHGITLQRRSESYSEDKCDMDSTNMDS NLKPEEVVHKEKRRTKSLLEEKLVLKSKSKTQG KQVKVVETELQEGATKQATTPKPDKEKNTEEND SEKQRKSKVEDKPFEETGVEPVLETASSSAHSTQ KDSSHRAKLPLAKEKYKSDKDSTSTRLERKLSD GHKSRSLKHSSKDIKKKDENKSDDKDGKEVDSS HEKARGNSSLMEKKLSRRLCENRRGSLSQEMAK GEEKLAANTLSTPSGSSLQRPKKSGDMTLIPEQEP MEIDSEPGVENVFEVSKTQDNRNNNSHQDIDSEN

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				MKQKTSATVQKDELRTCTADSKATAPAYKPGR GTGVNSNSEKHADHRSTLTKKMHIQSAVSKMNP GEKEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPSLSS VTVVPLRESYDPDVIPLFDKRTVLEGSTASTSPAD HSALPNQSLTVRESEVLKTSDSKEGGEGFTVDTP AKASITSKRHIPEAHQATLLDGKQGKVIMPLGSK LTGVIVENENITKEGGLVDMAKKENDLNAEPNL KQTIKATVENGKKDGIAVDHVVGLNTEKYAETV KLKHKRSPGKVKDISIDVERRNENSEVDTSAGSG SAPSVLHQRNGQTEDVATGPRRAEKTSVATSTE GKDKDVTLSPVKAGPATTTSSETRQSEVALPCTS IEADEGLIIGTHSRNNPLHVGAEASECTVFAAAEE GGAVVTEGFAESETFLTSTKEGESGECAVAESED RAADLLAVHAVKIEANVNSVVTEEKDDAVTSAG SEEKCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEGTV TCTGAEGRSDNFVICSVTGAGPREERMVTGAGV VLGDNDAPPGTSASQEGDGSVNDGTEGESAVTS TGITEDGEGPASCTGSEDSSEGFAISSESEENGESA MDSTVAKEGTNVPLVAAGPCDDEGIVTSTGAKE EDEEGEDVVTSTGRGNEIGHASTCTGLGEESEGV LICESAEGDSQIGTVVEHVEAEAGAAIMNANENN VDSMSGTEKGSKDTDICSSAKGIVESSVTSAVSG KDEVTPVPGGCEGPMTSAASDQSDSQLEKVEDT TISTGLVGGSYDVLVSGEVPECEVAHTSPSEKED EDIITSVENEECDGLMATTASGDITNQNSLAGGK NQGKVLIISTSTTNDYTPQVSAITDVEGGLSDALR TEENMEGTRVTTEEFEAPMPSAVSGDDSQLTASR SEEKDECAMISTSIGEEFELPISSATTIKCAESLQP VAAAVEERATGPVLISTADFEGPMPSAPPEAESP LASTSKEEKDECALISTSIAEECEASVSGVVVESE NERAGTVMEEKDGSGIISTSSVEDCEGPVSSAVP QEEGDPSVTPAEEMGDTAMISTSTSEGCEAVMIG AVLQDEDRLTITRVEDLSDAAIISTSTAECMPISA SIDRHEENQLTADNPEGNGDLSATEVSKHKVPM PSLIAENNCRCPGPVRGGKEPGPVLAVSTEEGHN GPSVHKPSAGQGHPSAVCAEKEEKHGKECPEIGP FAGRGQKESTLHLINAEEKNVLLNSLQKEDKSPE TGTAGGSSTASYSAGRGLEGNANSPAHLRGPEQ TSGQTAKDSSVSSIRYLAAVNTGAIKADDMPPVQ GTVAEHSFLPAEQQGSEDNLKTSTTKCITGQESKI APSHTMIPPATYSVALLAPKCEQDLTIKNDYSGK WTDQASAEKTGDDNSTRKSFPEEGDIMVTVSSE ENVCDIGNEESPLNVLGGLKLKANLKMEAYVPS EEEKNGEILAPPESLCGGKPSGIAELQREPLLVNE SLNVENSGFRTNEEIHSESYNKGEISSGRKDNAE AISGHSVEADPKEVEEEERHMPKRKRKQHYLSSE DEPDDNPDVLDSRIETAQRQCPETEPHATKEENS RDLEELPKTSSETNSTTSRVMEEKDEYSSSETTGE KPEQNDDDTIKSQE
3916	A	2	773	GPFGVLWPSAKPGPVTAVEARPPDASDPEGLRG GSPAPLLAPGPLDPSGRLHPAVSMMSYLKQPPYG MNGLGLAGPAMDLLHPSVGYPATPRKQRRERTT FTRSQLDVLEALFAKTRYPDIFMREEVALKINLPE

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SRVQVWFKNRRAKCRQQQQSGSGTKSRPAKKK SSPVRESSGSESSGQFTPPAVSSSASSSSSASSSSA NPAAAAAAGLVVAKLPCPLHIFSLCVFIEENRLV SGSWARDIRSVEETDKSGYR
3917	A	2	776	RNIPGRRFRPPGLRRLLKGPHMPREPRGYRTRVP ALRELVPSSHAGSGASEHCQNNRQGSRQHRASR NVQAGGALAPPRHLCGLCSRLHFLKPDLSVRAA PSRAGASVMALRKELLKSIWYAFTALDVEKSGK VSKSQLRVLSHNLYTVLHIPHDPVALEEHFRDDD DGPVSSQGYMPYLNKYILDKVEEGAFVKEHFDE LCWTLTAKKNYRADSNGNSMLSNQDAFRLWCL FNFLSEDKYPLIMDPDEGEYLLKRYS
3918	A	10	318	WQDLVCLGGSRAQEQKPLQQLWNAILLVAMLL CTGLVVQAQRQASRQSQRELGGQVDLFKRRVV RRLASLKTRRCRLSRAAQGLPDPGAETCAVCLD YFCNKQ
3919	A	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPKKT RAMRRRLNMHEENLKTKKQHRKERLYPLRKYA AKA
3920	A	1	654	RCCRSFVAPLQEKVVFGLFFLGAILCLSFSWLFHT VYCHSEGVSRLFSKLDYSGIALLIMGSFVPWLYY SFYCNPQPCFIYLIVICVLGIAAIIVSQWDMFATPQ YRGVRAGVFLGLGLSGIIPTLHYVISEGFLKAATI GQIGWLMLMASLYITGAALYAARIPERFFPGKCD IWFHSHQLFHIFVVAGAFVHFHGVSNLQEFRFMI GGGCSEEDAL
3921	A	1587	452	LERDGCGGEEGGSVRSGAGPDSDPRGASSPPAG HRGTAASPRPVAAPSRTPAPPHTRARASPGLPSG PAWRRVQWFSRVSGQVSTLMKATVLMRQPGRV QEIVGALRKGGGDRLQVISDFDMTLSRFAYNGK RCPSSYNILDNSKIISEECRKELTALLHHYYPIEID PHRTVKEKLPHMVEWWTKAHNLLCQQKIQKFQI AQVVRESNAMLREGYKTFFNTLYHNNIPLFIFSA GIGDILEEIIRQMKVFHPNIHIVSNYMDFNEDGFL QGFKGQLIHTYNKNSSACENCGYFQQLEGKTNV ILLGDSIGDLTMADGVPGVQNILKIGFLNDKVEE RRERYMDSYDIVLEKDETLDVVNGLLQHILCQG VQLEMQGP
3922	A	2	164	GKIYQRAFGGHSLKFGKGVQAHGCCCVADRTG HSILHTSYGRERPAPVHLRQDT
3923	A	2	3258	EHATHAYAKLGTRRRHREVTVFVPTWQLKKNR RVRESHFLTKLHSLKMLSITPSQLENGKKITTYD YRFMVKLAEETDGIIVTNEQIHILMNSSKKLMVK DRLLPFTFAGNLFMVPDDPLGRDGPTLDEFLKKP NRLDTDIGNFLKVWKTLPPSSASVTELSDDADSG PLESLPNMEEVREEKEERQDEEQRQGQGTQKAA EEDDLDSSLASVFRVECPSLSEEILRCLSLHDPPD GALDIDLLPGAASPYLGIPWDGKAPCQQVLAHL AQLTIPSNFTALSFFMGFMDSHRDAIPDYEALVG PLHSLLKQKPDWQWDQEHEEAFLALKRALVSAL CLMAPNSQLPFRLEVTVSHVALTAILHQEHSGRK HPIAYTSKPLLPDEESQGPQSGGDSPYAVAWALK HFSRCIGDTPVVLDLSYASRTTADPEVREGRRVS KAWLIRWSLLVQDKGKRALELALLQGLLGENRL LTPAASMPRFFQVLPPFSDLSTFVCIHMSGYCFYR

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EDEWCAGFGLYVLSPTSPPVSLSFSCSPYTPTYAHLAAVACGLERFGQSPLPVVFLTHCNWIFSLLWELLPLWRARGFLSSDGAPLPHPSLLSYIISLTSGLSSLPFIYRTSYRGSLFAVTVDTLAKQGAQGGGQWWSLPKDVPAPTVSPHAMGKRPNLLALQLSDSTLADIIARLQAGQKLSGSSPFSSAFNSLSLDKESGLLMFKGDKKPRVWVVPTQLRRDLIFSVHDIPLGAHQRPEETYKKLRLLGWWPGMQEHVKDYCRSCLFCIPRNLIGSELKVIESPWPLRSTAPWSNLQIEVVGPVTISEEGHKHVLIVADPNTRWVEAFPLKPYTHTAVAQVLLQHVFARWGVPVRLEAAQGPQFARHVLVSCGLALGAQVASLSRDLQFFCLTSSGAYWEFKRALKEFIFLHGKKWAASLPLLHLAFRASSTDATPFKVLTGGESRLTEPLWWEMSSANIEGLKMDVFLLQLVGELLELHWRVADKASEKAENRRFKRESQEKEWNVGDQVLLLSLPRNGSSAKWVGPFYIGDRLSLSLYRIWGFPTPEKLGCIYPSSLMKAFAKSGTPLSFKVLEQ
3924	A	1	1826	MGSVTVRYFCYGCLFTSATWTVLLFVYFNFSEVTQPLKNVPVKGSGPHGPSPKKFYPRFTRGPSRVLEPQFKANKIDDVIDSRVEDPEEGHLKFSSELGMIFNERDQELRDLGYQKHAFNMLISDRLGYHRDVPDTRNAACKEKFYPPDLPAASVVICFYNEAFSALLRTVHSVIDRTPAHLLHEIILVDDDSDFDDLKGELDEYVQKYLPGKIKVIRNTKREGLIRGRMIGAAHATGEVLVFLDSHCEVNVMWLQPLLAAIREDRHTVGCPVIDIISADTLAYSSSPVVRGGFNWGLHFKWDLVPLSELGRAEGATAPIKSPTMAGGLFAMNRQYFHELGQYDSGMDIWGGENLEISFRIWMCGGKLFIIPCSRVGHIFRKRRPYGSPEGQDTMTHNSLRLAHVWLDEYKEQYFSLRPDLKTKSYGNISERVELRKKLGCKSFKWYLDNVYPEMQISGSHAKPQQPIFVNRGPKRPKVLQRGRLYHLQTNKCLVAQGRPSQKGGLVVLKACDYSDPNQIWIYNEEHELVLNSLLCLDMSETRSSDPPRLMKCHGSGGSQQWTFGKNNRLYQVSVGQCLRAVDPLGQKGSVAMAICDGSSSQQWHLEG
3925	A	5386	2897	VRWNSKTECYLSIQTQENFPANLNELVNCIVISSLVTTQRKLKAMSLLGSRNQLARAVLNPNPMDFCTKDLLTTTSERIIAYLRDFNEDQKKAIETAYAMVKHSPSVAKICLIHGPPGTGKSKTIVGLLYRLLTENQRKGHSDENSNAKIKQNRVLVCAPSNAAVDELMKKIILEFKEKCKDKKNPLGNCGDINLVRLGPEKSINSEVLKFSLDSQVNHRMKKELPSHVQAMHKRKEFLDYQLDELSRQRALCRGGREIQRQELDENISKVSKERQELASKIKEVQGRPQKTQSIIILESHIICCTLSTSGGLLLESAFRGQGGVPFSCVIVDEAGQSCEIETLTPLIHRCNKLILVGDPKQLPPTVISMKAQEYGYDQSMMARFCRLLEENVEHNMISRLPILQLTVQYRMHPDICLFPSNYVYNRNLKTNRQTEAIRCSSDWPFQPYLVFDVGDGSERRDNDSYINVQEIKLVMEIIKLIKDKRKDVSFRNIGIITHYKAQKTMIQKDLDKEFDRKGPAEVDTVDAFQGRQKDCVIVTCVRANSIQGSIGFLASLQRLNVTITRAKYSLFILGHLRTLMENQHWNQLIQDAQKRGAIIKTCDKNYRHDAV

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				KILKLKPVLQRSLTHPPTIAPEGSRPQGGLPSSKLDSGFAKTSVAASLYHTPSDSKEITLTVTSKDPERPPVHDQLQDPRLLKRMGIEVKGGIFLWDPQPSSPQHPGATPPTGEPGFPVVHQDLSHVQQPAAVVAALSSHKPPVRGEPPAASPEASTCQSKCDDPEEELCHRREARAFSEGEQEKCGSETHHTRRNSRWDKRTLEQEDSSSKKRKLL
3926	A	99	284	MPREDRATWKSNYFLKIIQLLDDYPKRFIVGANNVGSKQMQQIRMSLRGKAVVLMGKNTMMR
3927	A	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGENWIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRYCVIIHPMSCFSIHKTRCAVVACAVVWIISLVAVIPMTFLITSTNRTNRSACLDLTSSDELNTIKWYNLILTA\LLCLPLVIVTLCYTTIIHTLTHGHAN\DSCLKQKARRLTILLL
3928	A	1	1516	GEEAVGGGAEGGGFGVGAQGRAGGRGVEAGRMRLSKTLVDMDMADYSAALDPAYTTLEFENVQVLTMGNDTSPSEGTNLNAPNSLGVSALCAICGDRATGKHYGASSCDGCKGFFRRSVRKNHMYSCRFSRQCVVDKDKRNQCRYCRLKKCFRAGMKKEAVQNERDRISTRRSSYEDSSLPSINALLQAEVLSRQITSPVSGINGDIRAKKIASIADVCESMKEQLLVLVEWAKYIPGFCELPLDDQGALLRAHAGEHLLLGATKRSMVFKDVLLLGNDYIVPRHCPELAEMSRVSIRILDELVLPFQELQIDDNEYAYLKAIIFFDPDAKGLSDPGKIKRLRSQVQVSLEDYINDRQYDSRGRFGELLLLLPTLQSITWQMIEQIQFIKLFGMAKIDNLLQEMLLGGSPSDAPHAHHPLHPHLMQEHMGTNVIVANTMPTHLSNGQMCEWPRPRGQAATPETPQPSPPGASGSEPYKLLPGAVATIVKPLSAIPQPTITKQEVI
3929	A	1	2782	RVLSLESPLEKDPRVLGAQSVPRGRALKGLSPLGLDSAFRLFPDPRAGPWNTAVLSSGMEPETALWGPDLQGPEQSPNDAHRGAESENEEESPRQESSGEEIIMGDPAQSPESKDSTEMSLERSSQDPSVPQNPPTPLGHSNPLDHQIPLDPPAPEVVPTPSDWTKACEASWQWGALTTWNSPPVVPANEPSLRELVQGRPAGAEKPYICNECGKSFSQWSKLLRHQRIHTGERPNTCSECGKSFTQSSHLVQHQRTHTGEKPYKCPDCGKCFSWSSNLVQHQRTHTGEKPYKCTECEKAFTQSTNLIKHQRSHTGEKPYKCGECRRAFYRSSDLIQHQATHTGEKPYKCPECGKRFGQNHNLLKHQKIHAGEKPYRCTECGKSFIQSSELTQHQRTHTGEKPYECLECGKSFGHSSTLIKHQRTHLREDPFKCPVCGKTFTLSATLLRHQRTHTGERPYKCPECGKSFSVSSNLINHQRIHRGERPYICADCGKSFIMSSTLIRHQRIHTGEKPYKCSDCGKSFIRSSHLIQHRRTHTGEKPYKCPECGKSFSQSSNLITHVRTHMDENLFVCSDCGKAFLEAHELEQHRVIHERGKTPARRAQGDSLLGLGDPSLLTPPPGAKPHKCLVCGKGFNDEGIFMQHQRIHIGENPYKNADGLIAHAAPKPPQLRSPRLPFRGNSYPGAAEGRAEAPGQPLKPPEGQEGFSQRRGLLSSKTYICSHCGESFLDRSVLLQHQLTHGNEKPFLFPDYRIGLGEGAGPSPFLSGKPFKCPECKQSFGLSSELLLHQKVHAGGKSSHKSPELGKSSSVLL

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				EHLRSPLGARPYRCSDCRASFLDRVALTRHQETH TQEKPPNPEDPPPEAVTLSTDQEGEGETPTPTESS SHGEGQNPKTLVEEKPYLCPECGAGFTEVAALLL HRSCHPGVSL
3930	A	513	273	KTQETHIYISEHIFFPFLQGFGNLPICMAKTDLSLS HQPDKKGVPSDFILPISDVRASIGAGFIYPLVGTG SRESPLWL
3931	A	16	305	KRRDFLSCWPAFTVLGEARGDQVDWSKLYRDT GLVKMSRKPRASSPFSNNHPSTPKRRGRGKHPLI PGPEALSKFPRQPIREKGPVKEVPGTKGSP
3932	A	16	305	KRRDFLSCWPAFTVLGEARGDQVDWSKLYRDT GLVKMSRKPRASSPFSNNHPSTPKRRGRGKHPLI PGPEALSKFPRQPIREKGPVKEVPGTKGSP
3933	A	1	1546	STHASEHWDSALQLAKHLAPDQIPFISKEYAIQLE FAGDYVNALAHYEKGITGDNKEHDEACLAGVA QMSIRMGDIRRGVNQALKHPSRVLKRDCGAILE NMKQFSEAAQLYEKGLYYDKAASVYIRSKNWA KVGDLLPHVSSPKIHLQYAKAKEADGRYKEAVV AYENAKQWQSVIRIYLDHLNNPEKAVNIVRETQ SLDGAKMVARFFLQLGDYGSAIQFLVMSKCNNE AFTLAQQHNKMEIYADIIGSEDTTNEDYQSIALY FEGEKRYLQAGKFFLLCGQYSRALKHFLKCPSSE DNVAIEMAIETVGQAKDELLTNQLIDHLLGEND GMPKDAKYLFRLYMALKQYREAAQTAIIIAREE QSAGNYRNAHDVLFSMYAELKSQKIKIPSEMAT NLMILHSYILVKIHVKNGDHMKGARMLIRVANN ISKFPSHIVPILTSTVIECHRAGLKNSAFSFAAML MRPEYRSKIDAKYKKKIEGMVRRPDISEIEEATTP CPFCKFLLPESELL
3934	A	334	1268	PTRRPILPLTSPKAISVPSPLQGKQHTLVKSCLSVS GIGGFLVSLSSRMKLQTLAVSVTALKFWSAYVP CQTQDRDALRLTLEQIDLIRRMCASYSELELVTS AKALNDTQKLACLIGVEGGHSLDNSLSILRTFYM LGVRYLTLTHTCNTPWAESSAKGVHSFYNNISGL TDFGEKVVAEMNRLGMMVDLSHVSDAVARRAL EVSQAPVIFSHSAARGVCNSARNVPDDILQLLEE ERWAFVMVSLFHGELIQWQPIRPMCSTVADHFD HIKAV\IGSKFIGIGGDYDGAGKYRKKTTCKAPW RTSSRMSS
3935	A	1	883	HETTPAVVQSVLLERGWNKFDKQEQNAEDWNL YWRTSSFRMTEHNSVKPWQQLNHHPGTTKLTR KDCLAKHLKHMRRMYGTSLYQFIPLTFVMPNDY TKFVAEYFQERQMLGTKHSYWICKPAELSRGRG ILIFSDFKDFIFDDMYIVQKYISNPLLIGRYKCDLR IYVCVTGFKPLTIYVYQEGLVRFATEKFDLSNLQ NNYAHLTNSSINKSGASYEKIKEVIGHGCKWTLS RFFSYLRSWDVDDLLLWKKIHRMVILTILAIAPS VPFAANCFELFGFDILIDDNEFHRTG
3936	A	203	441	HLAHSLGPLPKHYQYCVRYLYYQVTKDVIKEFA DDGVKYLELRSTPRRENATGMTKKTYVESILEGI KQSKQENLDIDV

TABLE 7

SEQ ID NO:	Position of end of Signal in Amino Acid Sequence	MaxS (MAXIMUM SCORE)	MeanS (Mean Score)
1	19	0.930	0.680
2	24	0.964	0.863
3	21	0.990	0.901
4	19	0.981	0.942
5	22	0.991	0.928
6	21	0.956	0.843
8	22	0.913	0.718
9	17	0.997	0.969
11	19	0.930	0.680
13	36	0.983	0.863
14	28	0.935	0.839
15	21	0.997	0.955
16	16	0.983	0.944
17	18	0.989	0.884
19	49	0.996	0.719
20	28	0.972	0.920
21	23	0.954	0.905
22	46	0.955	0.568
23	26	0.942	0.654
24	19	0.979	0.941
25	34	0.884	0.565
26	33	0.934	0.584
27	17	0.975	0.914
28	18	0.980	0.934
29	23	0.928	0.718
30	26	0.978	0.885
32	20	0.946	0.719
33	29	0.933	0.671
35	25	0.996	0.920
36	26	0.903	0.579
40	19	0.981	0.942
47	25	0.971	0.909
53	22	0.991	0.928
55	24	0.960	0.808
60	19	0.986	0.967
78	22	0.913	0.718
86	20	0.883	0.555
87	24	0.982	0.889
88	17	0.997	0.969
115	19	0.930	0.680
134	36	0.983	0.863
136	17	0.913	0.696
137	19	0.958	0.905
140	28	0.935	0.839
143	32	0.914	0.740
153	21	0.997	0.955
154	25	0.913	0.583
155	29	0.972	0.857
169	30	0.977	0.817
170	30	0.977	0.819
171	30	0.977	0.819
175	47	0.926	0.606
176	30	0.968	0.872
177	22	0.957	0.791
192	43	0.930	0.678

SEQ ID NO:	Position of end of Signal in Amino Acid Sequence	MaxS (MAXIMUM SCORE)	MeanS (Mean Score)
195	19	0.956	0.860
202	21	0.982	0.871
203	24	0.957	0.870
207	23	0.954	0.905
224	46	0.955	0.568
225	26	0.942	0.654
228	45	0.961	0.839
231	28	0.994	0.937
232	28	0.993	0.896
234	19	0.979	0.942
235	19	0.979	0.941
238	20	0.987	0.943
244	23	0.929	0.683
250	34	0.884	0.565
256	33	0.934	0.584
258	25	0.934	0.729
259	22	0.969	0.871
264	19	0.952	0.753
265	17	0.975	0.914
266	17	0.975	0.914
271	23	0.974	0.884
274	13	0.971	0.834
275	18	0.980	0.934
278	32	0.958	0.668
280	24	0.966	0.881
281	24	0.966	0.881
286	23	0.928	0.718
291	35	0.991	0.824
293	27	0.956	0.806
294	23	0.952	0.827
301	26	0.978	0.885
316	20	0.946	0.719
320	28	0.978	0.726
327	29	0.933	0.671
331	48	0.903	0.571
345	25	0.996	0.920
349	26	0.903	0.579
351	24	0.951	0.876
352	18	0.944	0.716
353	32	0.992	0.854
354	27	0.945	0.817
355	16	0.922	0.716
356	13	0.959	0.818
357	23	0.986	0.878
358	19	0.904	0.671
359	16	0.988	0.951
360	15	0.981	0.938
361	18	0.944	0.716
362	21	0.984	0.869
363	40	0.979	0.813
364	18	0.883	0.693
365	22	0.962	0.908
366	22	0.961	0.827
367	44	0.941	0.624
368	20	0.952	0.791
369	22	0.949	0.840
370	28	0.957	0.682

SEQ ID NO:	Position of end of Signal in Amino Acid Sequence	MaxS (MAXIMUM SCORE)	MeanS (Mean Score)
372	28	0.974	0.894
373	19	0.972	0.947
374	29	0.968	0.785
375	19	0.949	0.897
377	23	0.962	0.910
378	31	0.974	0.895
379	26	0.969	0.939
380	27	0.945	0.817
383	27	0.945	0.817
384	25	0.992	0.877
385	32	0.983	0.825
386	44	0.924	0.564
387	26	0.971	0.894
388	19	0.989	0.862
389	24	0.990	0.947
390	34	0.942	0.635
391	16	0.922	0.716
394	19	0.987	0.970
398	36	0.992	0.866
404	13	0.959	0.818
417	23	0.986	0.878
421	19	0.904	0.671
425	28	0.971	0.717
431	16	0.988	0.951
452	18	0.944	0.716
459	21	0.991	0.902
468	21	0.984	0.869
478	40	0.979	0.813
486	18	0.883	0.693
499	22	0.962	0.908
501	19	0.962	0.877
514	44	0.941	0.624
529	20	0.952	0.791
533	39	0.914	0.719
548	28	0.957	0.682
561	28	0.974	0.894
562	28	0.974	0.893
564	18	0.949	0.806
576	19	0.972	0.947
584	29	0.968	0.785
585	28	0.973	0.810
591	19	0.949	0.897
592	24	0.991	0.954
594	20	0.985	0.959
595	20	0.985	0.959
612	23	0.962	0.910
619	31	0.974	0.895
621	15	0.959	0.795
633	26	0.969	0.939
640	20	0.949	0.842
645	25	0.911	0.759
684	25	0.992	0.877
691	32	0.983	0.825
698	44	0.924	0.564
700	19	0.982	0.941
710	26	0.971	0.894
714	23	0.965	0.907

SEQ ID NO:	Position of end of Signal in Amino Acid Sequence	MaxS (MAXIMUM SCORE)	MeanS (Mean Score)
718	19	0.989	0.862
725	21	0.976	0.851
728	33	0.961	0.895
734	25	0.963	0.660
741	34	0.942	0.635
744	19	0.959	0.924
747	16	0.922	0.716
756	26	0.973	0.864
767	22	0.986	0.943
768	27	0.916	0.758
769	19	0.987	0.970
770	22	0.981	0.933
771	34	0.993	0.893
773	20	0.968	0.939
774	21	0.971	0.945
778	22	0.986	0.943
779	32	0.973	0.846
781	23	0.950	0.857
785	27	0.916	0.758
786	27	0.916	0.758
788	22	0.981	0.933
793	22	0.986	0.803
794	39	0.892	0.654
797	27	0.965	0.847
810	22	0.981	0.933
823	34	0.993	0.893
825	17	0.962	0.778
837	20	0.968	0.939
844	25	0.984	0.951
845	17	0.919	0.706
846	21	0.971	0.945
847	21	0.971	0.945
890	22	0.986	0.943
893	24	0.971	0.865
894	24	0.971	0.865
896	32	0.973	0.846
899	31	0.982	0.817
922	15	0.882	0.706
924	21	0.975	0.948
925	21	0.927	0.661
933	20	0.967	0.906
960	20	0.967	0.906
967	38	0.970	0.784
968	47	0.970	0.557
972	36	0.945	0.775

TABLE 8

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
3955	A	235	1272	GPREVLAASSLADGSEEQVMAVALVRERDLSFPG VGDAVVNPTRWHLPAQPEMLYEGGEGRMETLK

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				DKTLQELEELQNDSEAIDQLALESPEVQDLQLERE MALATNRSLAERNLEFQGPLEISRSNLSDRYQELR KLVERCQEQKAKLEKFSSALQPGTLLDLLQVEGM KIEEESEAMAEKFLEGEVPLETFLENFSSMRMLSH LRRVRVEKLQEVVRKPRASQELAGDAPPPRSPPP V/PPSPPGNTPCG*RAAAATISHASLPFALQPIPQPA CGPHCPWSPATGPFPSSVPALLLQRASGPHLPGSP AWTQGCCGLLLVPTEEHAAPPYGFPPPPGPAWPG Y
3956	A	821	385	SICADRTERVGIFFYIPAGTTDEADVTHP*EGHSYL SNHAGIQRSSRP/SHYQGE/WHDNCFTADELQLLT YQLCHTYVRCTRSVSIPAPAYYAHLVAFRARYHL VDKEHDSAEGSHVSGQSNGRDPQALAKAVQIHQ DTLRTMYFA
3957	A	4621	240	ELISTFKLLLEKKRSEVMKMKKRYEVGLEKLDSA SSQVATMQMELEALHPQLKVASKEVDEMMIMIE KESVEVAKTEKIVKADETIANEQAMASKAIKDEC DADLAGALPILESALAALDTLTAQDITVVKSMKSP PAGVKLVMEAICILKGIKADKIPDPTGSGKKIEDF WGPAKRLLGDMRFLQSLHEYDKDNIPPAYMNIIR KNYIPNPDFVPEKIRNASTAAEGLCKWVIAMDSY DKVAKIVAPKKIKLAAAEGELKIAMDGLRKKQA ALKEVQDKLARLQDTLELNKQKKADLENQVDLC SKKLERAEQLIGGLGGEKTRWSHTALELGQLYIN LTGDILISSGVVAYLGAFTSTYRQNQTKEWTTLCK GRDIPCSDDCSLMGTLGEAVTIRTWNIAGLPSDSF SIDNGIIIMNARRWPLMIDPQSQANKWIKNMEKA NSLYVIKLSEPDYVRTLENCIQFGTPVLLENVGEE LDPILEPLLLKQTFKQGGSTCIRLGDSTIEYAPDFR FYITTKLRNPHYLPETSVKVTLLNFMITPEGMQDQ LLGIVVAQERPDLEEEKQALILQGAENKRQLKEIE DKILEVLSSSEGNILEDETAIKILSSSKALANEISQK QEVAEETEKKIDTTRMGYRPIAIHSSILFFSLADLA NIEPMYQYSLTWFINLFILSIENSEKSEILAKRLQIL KDHFTYSLYVNVCRSLFEKDKLLFSFCLTINLLLH ERAINKAEWRFLLTGGIGLDNPYANPCTWLPQKS WDEICRLDDLPAFKTIRREFMRLKDGWKKVYDSL EPHHEVFPEEWEDKANEFQRMLIIRCLRPDKVIPM LQEFIINRLGRAFIEPPPFDLAKAFGDSNCCAPLIFV LSPGADPMAALLKFADDQGYGGSKLSSLSLGQGQ GPIAMKMLEKAVKEGTWVVLQNCHLATSWMPT LEKVCEELSPESTHPDFRMWLTSYPSPNFPVSVLQ NGVKMTNEAPKGLRANIIRSYLMDPISDPEFFGSC KKPEEFKKLLYGLCFFHALVQERRKFGPLWWNIP YEFNETDLRISVQQLHMFLNQYEELPYEALRYMT GECNYGGRVTDDWDRRTLRSILNKFFNPELVENS DYKFDSSGIYFVPPSGDHKSYIEYTKTLPLTPAPEI FGMNANADITKDQSETQLLFDNILLTQSRSAGAG AKSSDEVVNEVASDILGKLPNNFDIEAAMRRYPT TYTQSMNTVLVQEMGRFNKLLKTIRDSCVNIQKA IKGLAVMSTDLEEVVSSILNVKIPEMWMGKSYPS LKPLGSYVNDFLARLKFLQQWYEVGPPPVFWLSG FFFTQAFLTGAQQNYARKYTIPIDLLGFDYEVMED KEYKHPPEDGVFIHGLFLDGASWNRKIKKLAESH PKILYDTVPVMWLKPCKRADIPKRPSYVAPLYKT

SEQ ID NO:	Method	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				SERRGVLSTTGHSTNFVIA\MTLPSDQPKEHWIGR GVALLCQLNS
3958	A	35	529	GADMAKSKNHTTHNQSRKWHRNVIKKPLSQRYK SLKGVDPKFLGNMCFTKKHKKKGLKKMQADSA KAVSTCAKAIEALVKPKEVKPKIPKGVSCELN*LA YIAYPKFWTCACACIAKGLRLCQPKAKAQDQTK AQVQIKAQAAAPASVPTQAPKGAQAPTKASG
3959	A	1883	763	LLVLLLRTNLLIASSTRISRATLTCSPPGIPVDPRVR PRVRSHLVMYLGITTGSLHKAVVSGDSSAHLVEEI QLFPDPEPVRNLQLAPTQGAVFVGFSGGVWRVPR ANCSVYESCVDCVLARDPHCAWDPESRTCCLLSA PNLNSWKQDMERGNPEWACASGPMSRSLRPQSR PQIIKEVLAVPNSILELPCPHLSALASYYWSHGPAA VPEASSTVYNGSLLLIVQDGVGGLYQCWATENGF SYPVISYWVDSQDQTLALDPELAGIPREHVKVPLT RVSGGAALAAQQSYWPHFVTVTVLFALVLSGALI ILVASPLRALRARGKVQGCETLRPGEKAPLSREQH LQSPKECRTSASDVDADNNCLGTEVA
3960	A	1	481	SYAAPSLFVKSLYWALAFMAVLLAVSGVVIVVLA SRAGARCQQCPPGWVLSEEHCYYFSAEAQAWEA SQAFCSAYHATLPLLSHTQDFLGRYPVSRHSWVG AWRGPQGWHWIDEAPLPPQLLPEDGEDNLDINCG ALEEGTLVAANCSTPRPWVCAKGTQ

TABLE 9

SEQ ID NO:	Accession Number	Species	Description	Smith Waterman Score	% Idenity
3937	Y27700	Homo sapiens	Human secreted protein encoded by gene No. 12.	193	25
3938	AF093097	Homo sapiens	putative RNA-binding protein Q99	3881	84
3939	AB012308	Anthocidaris crassispina	B2HC	4169	74
3940	U10248	Homo sapiens	ribosomal protein L29	787	95
3941	Y99418	Homo sapiens	Human PRO1317 (UNQ783) amino acid sequence SEQ ID NO:277.	4031	100
3942	AL023516	Gallus gallus	B locus C type Lectin	198	35

TABLE 10

SEQ ID NO:	Accession No.	Description	Results*
3937	PR00049	WILM'S TUMOUR PROTEIN SIGNATURE	PR00049D 0.00 9.168e-11 209-224
3942	BL00615	C-type lectin domain proteins.	BL00615A 16.68 6.400e-11 37-55

* Results Include in order: accession number subtype; raw score; p-value; position of signature in amino acid sequence

TABLE 11

SEQ ID NO:	PFAM Name	Description	P-Value	PFAM Score
3938	Piwi	Piwi domain	2.6e-150	512.7
3940	Ribosomal_L29e	Ribosomal L29e protein family	2.3e-19	77.8
3941	Sema	Sema domain	4e-181	615.1
3942	lectin_c	Lectin C-type domain	0.086	-7.1

TABLE 12

SEQ ID NO:	Position of end of Signal in Amino Acid Sequence	MaxS (Maximum Score)	Means (Mean Score)
3941	31	0.985	0.926
3942	21	0.974	0.894

TABLE 13

SEQ ID NO: of full length nucleotide sequence	SEQ ID NO: of full length peptide sequence	SEQ ID NO: of contig nucleotide sequence	SEQ ID NO: of contig peptide sequence	Priority Docket number corresponding SEQ ID NO: in priority application	SEQ ID NO: in USSN 09/496,914
3937	3943	3949	3955	787CIP2G_1	787_3587
3938	3944	3950	3956	787CIP2G_2	787_3813
3939	3945	3951	3957	787CIP2G_3	787_4462
3940	3946	3952	3958	787CIP2G_4	787_4887
3941	3947	3953	3959	787CIP2G_5	787_5794
3942	3948	3954	3960	787CIP2G_6	787_8743

TABLE 14

TISSUE ORIGIN	LIBRARY/ RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS:
adult brain	GIBCO	ABD003	3940
adult brain	Clontech	ABR006	3940
adult brain	Invitrogen	ABR014	3940
cultured preadipocytes	Strategene	ADP001	3937
adult heart	GIBCO	AHR001	3940
adult kidney	GIBCO	AKD001	3940
adult lung	GIBCO	ALG001	3940
young liver	GIBCO	ALV001	3940
adult ovary	Invitrogen	AOV001	3938, 3940-3941
adult spleen	GIBCO	ASP001	3940-3941
testis	GIBCO	ATS001	3940
bone marrow	Clontech	BMD001	3938, 3940
bone marrow	Clontech	BMD004	3940
adult cervix	BioChain	CVX001	3940
endothelial cells	Strategene	EDT001	3940
fetal brain	Clontech	FBR006	3940
fetal brain	Invitrogen	FBT002	3940-3941
fetal heart	Invitrogen	FHR001	3940
fetal kidney	Clontech	FKD001	3940
fetal kidney	Clontech	FKD002	3940

TISSUE ORIGIN	LIBRARY/ RNA SOURCE	HYSEQ LIBRARY NAME	SEQ ID NOS:
fetal liver-spleen	Columbia University	FLS001	3937, 3940
fetal liver-spleen	Columbia University	FLS002	3938, 3941
fetal liver-spleen	Columbia University	FLS003	3940
fetal liver	Clontech	FLV004	3940
fetal skin	Invitrogen	FSK001	3940-3942
fetal spleen	BioChain	FSP001	3940
fetal brain	GIBCO	HFB001	3937, 3940-3941
infant brain	Columbia University	IB2002	3937, 3939, 3941
leukocyte	GIBCO	LUC001	3940-3941
leukocyte	Clontech	LUC003	3940-3941
melanoma from cell line ATCC #CRL 1424	Clontech	MEL004	3940
mammary gland	Invitrogen	MMG001	3937, 3940-3941
neuronal cells	Strategene	NTU001	3937, 3942
prostate	Clontech	PRT001	3938
rectum	Invitrogen	REC001	3940
salivary gland	Clontech	SALs03	3941
small intestine	Clontech	SIN001	3940
skeletal muscle	Clontech	SKM001	3940
spinal cord	Clontech	SPC001	3940
thymus	Clontech	THMc02	3938
thyroid gland	Clontech	THR001	3942
uterus	Clontech	UTR001	3940

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954, a full length protein coding portion of SEQ ID NO:1-984, 1969-2952, 3937-3942 or 3949-3954, a mature protein coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, an active domain coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, and complementary sequences thereof.

2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.

3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.

4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.

5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.

6. A vector comprising the polynucleotide of claim 1.

7. An expression vector comprising the polynucleotide of claim 1.

8. A host cell genetically engineered to comprise the polynucleotide of claim 1.

9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.

10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:

(a) a polypeptide encoded by any one of the polynucleotides of claim 1; and

(b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954.

11. A composition comprising the polypeptide of claim 10 and a carrier.

12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.

14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;

b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and

c) detecting said product and thereby the polynucleotide of claim 1 in the sample.

15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

16. A method for detecting the polypeptide of claim 10 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and

b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.

17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.

19. A method of producing the polypeptide of claim 10, comprising,

a) culturing a host cell comprising a polynucleotide sequence selected fromm the group consisting of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, a mature protein coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, an active domain coding portion of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954, under conditions sufficient to express the polypeptide in said cell; and

b) isolating the polypeptide from the cell culture or cells of step (a).

20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of any one of the polypeptides SEQ ID NO: 985-1968, 2953-3936, 3943-3948 or 3955-3960, the mature protein portion thereof, or the active domain thereof.

21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.

22. A collection of polynucleotides, wherein the collection comprising the sequence information of at least one of SEQ ID NO: 1-984, 1969-2952, 3937-3942 or 3949-3954.

23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.

24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.

25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.

26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.